

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 16:08:09 ; Search time 76 Seconds  
(without alignments)  
1002.886 Million cell updates/sec

Title: US-09-943-075a-2  
Perfect score: 3061  
Sequence: 1 MESLCGVLFLLAAGLPLQ.....PFSRGDRKDPQLQDKPWWL 572

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3040	99.3	572	19 AAW38335	Rat kidney injury
2	2700	88.2	574	21 AAB19119	Polypeptide isolat
3	2163	70.7	560	18 AAW35382	Murine metastatic
4	2163	70.7	560	21 AAB11329	Human lung cancer-
5	2163	70.7	560	23 ABB74961	Human lung tumour
6	2162	70.6	572	23 ABB74961	Human PRO protein,
7	644.5	21.1	376	18 AAW35386	Murine metastatic
8	644.5	21.1	376	18 AAW35387	Murine metastatic
9	642	21.0	625	20 AAW31978	Mouse melanoma ant
10	627.5	20.5	661	16 AAW78646	Melanoma associate

11	627.5	20.5	661	20 AAY31977	Human melanoma ant
12	627.5	20.5	661	22 AAB47540	Human melanoma ant
13	627.5	20.5	661	22 AAB47540	Melanoma antigen c
14	627.5	20.5	661	23 AAU84803	Human gp100 Consen
15	627.5	20.5	661	23 AAU84803	Human melanoma ant
16	627.5	20.5	662	20 AAY42627	Human melanoma ant
17	626.5	20.5	661	16 AAR84854	MART-1 melanoma an
18	626.5	20.5	661	22 AAU28912	Melanoma antigen c
19	624.5	20.4	661	22 AAB47544	Human melanoma ant
20	622.5	20.3	661	16 AAR84855	MART-1 melanoma an
21	622.5	20.3	661	22 AAU28912	Modified tumour-as
22	622.5	20.3	661	22 AAB98206	Human gp100 prote
23	622.5	20.3	661	22 AAB97816	Modified gp100 pr
24	619.5	20.2	661	22 AAB47542	Human melanoma ant
25	619.5	20.2	661	22 AAB47546	Human melanoma ant
26	619.5	20.2	661	22 AAB47547	Human melanoma ant
27	616.5	20.1	661	22 AAB47545	Human melanoma ant
28	615.5	20.1	661	22 AAB47548	Human melanoma ant
29	614.5	20.1	661	22 AAB47541	Human melanoma ant
30	614.5	20.1	661	22 AAB47543	Human melanoma ant
31	602	19.7	668	18 AAW38164	Pmel17 encoded by
32	462.5	15.1	105	22 ABB28900	Peptide #1581 enco
33	462.5	15.1	105	22 ABB34075	Peptide #1581 enco
34	462.5	15.1	105	22 ABB19513	Protein #1512 enco
35	462.5	15.1	105	22 AAB54857	Human brain expres
36	462.5	15.1	105	22 AAM67239	Human bone marrow
37	462.5	15.1	105	22 AAM15084	Peptide #1518 enco
38	462.5	15.1	105	22 AAM27532	Peptide #1569 enco
39	462.5	15.1	105	22 AAM02820	Peptide #1502 enco
40	462.5	15.1	105	23 ABB36890	Human peptide enco
41	295	9.6	69	22 ABB43987	Peptide #11493 enc
42	295	9.6	69	22 ABB26896	Peptide #8895 enco
43	295	9.6	69	22 ABB27162	Protein #9161 enco
44	255	9.6	69	22 AAM64989	Human brain expres
45	255	9.6	69	22 AAM77712	Human bone marrow

## ALIGNMENTS

RESULT 1  
AAW38335  
ID AAW38335 standard; Protein; 572 AA.  
XX  
AC AAW38335;  
XX  
DT 21-MAY-1998 (first entry)  
XX  
DE Rat kidney injury related molecule (KIM).  
XX  
DE Kidney injury related molecule; KIM; rat; renal disease; injury;  
KW nephritis; tissue regeneration; therapy; monoclonal antibody.  
XX  
OS Rattus sp.  
XX  
PN WO9744460-A1.  
XX  
PD 27-NOV-1997.  
XX  
PF 23-MAY-1997; 97WO-US09303.  
PR 23-AUG-1996; 96US-0023442.  
PR 24-MAY-1996; 96US-0018228.  
XX  
XX (BIOJ ) EIOGEN INC.  
PA Bonventre JV, Cate RL, Hession CA, Ichimura T, Sanicola-Nadel M; Wei H;  
XX  
XX WPI; 1998-018514/02.  
DR N-PSDB; AAT96034.  
XX  
PT DNA encoding kidney injury related molecule - which is upregulated

PT in injured or regenerating tissue, useful to promote growth of new  
XX tissue and survival of damaged tissue  
PS Claim 9; Page 41-43; 68pp: English.  
XX  
CC This protein, designated kidney injury related molecule (KIM), is  
CC up-regulated in injured or regenerating tissue. Its amino acid  
CC sequence was deduced from a clone (see AAT96034) obtained from  
CC ischaemic adult rat kidneys. A 307-amino acid rat KIM (see  
CC AAW38334) and a human KIM (see AAW38336) are also claimed. Recombinant  
CC KIM polypeptides can be expressed in prokaryotic and eukaryotic  
CC host cells using a claimed process. Soluble variants fused to  
CC a toxin, imageable compound or radionuclide, and IgG fusion  
CC proteins are also claimed. KIM, or an agonist, can used to treat  
CC renal disease and to promote the growth of new tissue or the  
CC survival of damaged tissue, generally in conditions where the  
CC binding of specific ligand to KIM stimulates cell growth, maintains  
CC cellular differentiation or reduces apoptosis, e.g. in cases of  
CC renal failure, nephritis, kidney transplants, toxic or hypoxic  
CC injury. A monoclonal antibody specific for KIM can be used to  
CC treat renal disease, e.g. where binding of KIM to ligand results in  
CC neoplasia, loss of cellular function, susceptibility to apoptosis  
CC or promotion of inflammation, deliver imaging agents to KIM  
CC expressing cells in vivo or in vitro and measure KIM concentration  
CC by immunoassay. Damage/regeneration of renal cells can be determined  
CC by measuring KIM, particularly to diagnose or monitor the progress  
CC of disease or therapy. KIM-expressing tumour cells can be inhibited  
CC by treatment with a fusion protein comprising KIM ligand or MAB with  
CC a toxin or radionuclide, and tumour cells that express KIM ligand  
CC can be inhibited with similarly tagged KIM or anti-KIM ligand  
XX antibody.  
SQ Sequence 572 AA;  
Query Match 99.3%; Score 3040; DB 19; Length 572;  
Best Local Similarity 99.3%; Pred. No. 7.1e-272;  
Matches 568; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MESLCGVLLVLLAAGLPLQAARFRDVLGHEQYPDHMRNNQLRGWSSDENWDEQLYP 60  
DB 1 MESLCGVLLVLLAAGLPLQAARFRDVLGHEQYPDHMRNNQLRGWSSDENWDEQLYP 60  
QY 61 VRRGEGRWKDSWGGRRVQAALTSDSPALVGSNITFVNLVFPKCEKEDANGNIVYERN 120  
DB 61 VRRGEGRWKDSWGGRRVQAALTSDSPALVGSNITFVNLVFPKCEKEDANGNIVYERN 120  
QY 121 RSDLELASDPYVYNTWTGADDEWEDNTSGQHLRFPDCKPPRPHGRKKWFFVVFHTL 180  
DB 121 RSDLELASDPYVYNTWTGADDEWEDNTSGQHLRFPDCKPPRPHGRKKWFFVVFHTL 180  
QY 181 GOYFQKLGCSARVSVINVLTVGPQVMEVIVFRHGRAYIPISKVDVYVITDQIPFV 240  
DB 181 GOYFQKLGCSARVSVINVLTVGPQVMEVIVFRHGRAYIPISKVDVYVITDQIPFV 240  
QY 241 TWYOKNRNSDFFLRDLPIFDVLIHDPISHLNYSATISYKWNKPNGLFVSNNTLN 300  
DB 241 TWYOKNRNSDFFLRDLPIFDVLIHDPISHLNYSATISYKWNKPNGLFVSNNTLN 300  
QY 301 HTYVLNGTFNFNTVQTAVPGPCPTPSPSSSTSPSPASSPSPVLPSPSLMPTGYS 360  
DB 301 HTYVLNGTFNFNTVQTAVPGPCPTPSPSSSTSPSPASSPSPVLPSPSLMPTGYS 360  
QY 361 MELSDISNENCRINRYGYFRATITVDGILEVNIIOVADVPIPTLPDNLMDIFVTCCK 420  
DB 361 MELSDISNENCRINRYGYFRATITVDGILEVNIIOVADVPIPTLPDNLMDIFVTCCK 420  
QY 421 ATPTEACTIISDPTCQIAQNRVCSVAVDELCLLSVRFAFNGSGTYCVNFTIGDDASIAL 480  
DB 421 ATPTEACTIISDPTCQIAQNRVCSVAVDELCLLSVRFAFNGSGTYCVNFTIGDDASIAL 480  
QY 481 TSALISIPGKDLGSPFTVNGVLISIGCLAMFVTWVITLLYKHKTYKPIGNCNTRNVVK 540  
DB 481 TSALISIPGKDLGSPFTVNGVLISIGCLAMFVTWVITLLYKHKTYKPIGNCNTRNVVK 540

QY 541 KGLSVPLSHAKAPFSRGDRKDPDLLODKPWML 572  
DB 541 KGLSVPLSHAKAPFSRGDRKDPDLLODKPWML 572  
RESULT 2  
AAB19119  
ID AAB19119 standard; Protein; 574 AA.  
XX AAB19119;  
AC AAB19119;  
DT 19-FEB-2001 (first entry)  
XX Polypeptide isolated from lymph node stromal cells of fsn +/- mice.  
XX Lymph node stromal cell; fsn +/- mice; inflammatory disorder;  
KW immune system disorder; cancer; viral disorder; HIV infection;  
KW blood vessel growth; tumour necrosis factor disorder; arthritis;  
KW inflammatory bowel disease; fibroblast growth factor-mediated disorder;  
KW cardiac failure.  
XX Mus sp.  
PN WO200058463-A1.  
XX  
PD 05-OCT-2000.  
XX  
PF 18-FEB-2000; 2000WO-NZ000015.  
XX  
PR 25-MAR-1999; 99US-0276268.  
PR 26-AUG-1999; 99US-0383586.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;  
PI Murison JG;  
XX  
DR WPI; 2000-664924/64.  
DR N-PSDB; AAA96729.  
XX  
PT Polypeptide expressed in mammalian fsn +/- lymph node stromal cells,  
PT useful for modulating growth of blood cells, for treating inflammatory  
PT and tumour necrosis factor-mediated disorders, cancer and viral  
PT disorders  
XX  
PS Claim 1; Page 54-55; 75pp: English.  
XX  
CC The present sequence represents a polypeptide sequence which is  
CC isolated from lymph node stromal cells of fsn +/- mice. The  
CC polynucleotides and their polypeptides are useful for treating an  
CC inflammatory disorder, disorder of immune system and cancer selected  
CC from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a  
CC viral disorder, in particular HIV infection and for modulating the  
CC growth of blood vessels. The polypeptides are useful for treating a  
CC tumour necrosis factor (TNF) mediated disorder, such as those selected  
CC from arthritis, inflammatory bowel disease and cardiac failure and a  
CC fibroblast growth factor-mediated disorder. It is also useful in assays  
CC to determine biological activity, to raise antibodies, to isolate  
CC corresponding ligands or receptors, to quantify levels of protein or  
CC cognate corresponding ligand or receptors, as antiinflammatory agents,  
CC and in compositions for the treatment of skin, connective tissue and  
CC immune system diseases. The polynucleotide is useful as marker for  
CC tissue, as a chromosome marker or tags in the identification of a  
CC genetic disorder.  
SQ Sequence 574 AA;  
Query Match 88.2%; Score 2700; DB 21; Length 574;  
Best Local Similarity 88.8%; Pred. No. 1.9e-240;  
Matches 507; Conservative 20; Mismatches 42; Indels 2; Gaps 1;  
QY 1 MESLCGVLLVLLAAGLPLQAARFRDVLGHEQYPDHMRNNQLRGWSSDENWDEQLYP 60

Db 1 MESLCGVLFLLAAGLPLOAKRFDVLGHEQYPNHREHQLRGSSDENEDHLYP 60  
Qy 61 VWRGEGRWKSWEGGRVQAALSDSPALVGSNITFVNNLPVPRCKEDANGNIVYERN 120  
Db 61 VWRGEGRWKSWEGGRVQAALSDSPALVGSNITFVNNLPVPRCKEDANGNIVYERN 120  
Qy 121 RSDLELASDPVYVNTWTTGADDEWENTSOQHLEPPDCKPPRPHGRKKKNFVVFHTL 180  
Db 121 RNDLGLTSLDLHVYNTWTTGADDEWENTSOQHLEPPDCKPPRPHGRKKKNFVVFHTL 180  
Qy 181 GOYFQKLGOCARSINTVNLTVGPQVMEVIVPRRHGRAYIPISKVKDYVITDQIPFV 240  
Db 181 GOYFQKLGOCARSINTVNLTVGPQVMEVIVPRRHGRAYIPISKVKDYVITDQIPFV 240  
Qy 241 TWYQKNDNRNSDETFLRDLPIFFDVLHDPSPFLNYSALSYKWNFGDNTGLFVSNHPTLN 300  
Db 241 TMSQKNDNRNSDETFLRDLPIFFDVLHDPSPFLNYSALSYKWNFGDNTGLFVSNHPTLN 300  
Qy 301 HTYVLNGTFNENLTVOAVPGPCPSPTPS--PSSSTSPSPASSPTLSTPSPSLMPTGY 358  
Db 301 HTYVLNGTFNENLTVOAVPGPCPSPTPS--PSSSTSPSPASSPTLSTPSPSLMPTGY 358  
Qy 359 KSMELSDISNENCRINRYGYFRATITIVDGIENVIIQVADVPITLQPDNSLMDFIYTC 418  
Db 361 KSMELSDISNENCRINRYGYFRATITIVDGIENVIIQVADVPITLQPDNSLMDFIYTC 420  
Qy 419 KGATPTEACTIISDPTCOIAONRVCSVAVDELCLLSVRRAPNGSGTYCVNFTLGDDASL 478  
Db 421 KGATPTEACTIISDPTCOIAONRVCSVAVDELCLLSVRRAPNGSGTYCVNFTLGDDASL 480  
Qy 479 ALTSALISIPGKOLGSLRTVNGVLISIGCLAMFVTMTVILLYKKHKYKPIGNCNTRNV 538  
Db 481 ALTSTLISIPGKOLGSLRTVNGVLISIGCLAMFVTMTVILLYKKHKYKPIGNCNTRNV 540  
Qy 539 KGKGLSVFLSHAKAPFSGRDKDPLLODKP 569  
Db 541 KGKGLSVFLSHAKAPFSGRDKDPLLODKP 571

RESULT 3  
AAW35382  
ID AAW35382 standard; Protein; 560 AA.  
XX  
AC AAW35382;  
DT 26-FEB-1998 (first entry)  
XX Murine metastatic nucleic acid sequence product.  
DE  
DE Mouse; murine; tumour; cancer; metastatic sequence; detection;  
KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;  
KW hypertrophy; screening.  
XX  
OS Mus musculus.  
XX  
PN W09718454-A2.  
XX  
PD 22-MAY-1997.  
XX  
PF 15-NOV-1996; 96W0-US18567.  
XX  
PR 30-JAN-1996; 96US-0594031.  
XX  
PR 16-NOV-1995; 95US-0006838.  
XX  
PA (THOM/) THOMPSON T.  
XX  
PI Thompson T;  
XX  
DR WPI; 1997-289397/26.  
XX  
XX Identifying tumour metastatic sequences - by introducing transfected  
PT cells into host mammal and analysing primary and metastatic  
PT

PT sequences by differential display PCR  
XX Disclosure; Fig 12CI; 102pp; English.  
XX Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old  
CC mouse embryos. The UGS cells were infected with retroviruses,  
CC cultured and implanted under the renal capsule of mice.  
CC Reconstitutions were harvested 5 weeks later, when they showed  
CC signs of distress from the tumour burden. Metastatised tumours were  
CC isolated from a site outside the renal capsule. RNA was isolated  
CC from primary tumours and metastases, reverse transcribed and  
CC subjected to differential display PCR. The sequences were analysed  
CC to obtain metastatic sequences, e.g. the sequence encoding the  
CC present sequence. The method can be used to detect, diagnose and  
CC treat disorders related to metastasis, or treat malignant or  
CC non-malignant disorders, e.g. hyperplasia, dysplasia and  
CC hypertrophy. The metastatic sequence can be used to screen a  
CC biological sample for metastasis, and it or its expression product  
CC may also be used to treat a metastatic disorder.  
XX Sequence 560 AA;  
SQ  
Query Match 70.7%; Score 2163; DB 18; Length 560;  
Best Local Similarity 69.4%; Pred. No. 8.2e-191;  
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;  
Qy 1 MESLCGVLFLLAAGLPLOAKRFDVLGHEQYPNHREHQLRGSSDENEDHLYP 60  
Db 1 MCLXYFGLFLLAARLPDLAAKRFHDVLGNRPSAYMREHNLQNGSSDENEDHLYP 60  
Qy 61 VWRGEGRWKSWEGGRVQAALSDSPALVGSNITFVNNLPVPRCKEDANGNIVYERN 120  
Db 61 VWRGEGRWKSWEGGRVQAALSDSPALVGSNITFVNNLPVPRCKEDANGNIVYERN 120  
Qy 121 RSDLELASDPVYVNTWTTGADDEWENTSOQHLEPPDCKPPRPHGRKKKNFVVFHTL 180  
Db 121 RNEAGLSADPVYVNTWTTGADDEWENTSOQHLEPPDCKPPRPHGRKKKNFVVFHTL 180  
Qy 181 GOYFQKLGOCARSINTVNLTVGPQVMEVIVPRRHGRAYIPISKVKDYVITDQIPFV 240  
Db 181 GOYFQKLGOCARSINTVNLTVGPQVMEVIVPRRHGRAYIPISKVKDYVITDQIPFV 240  
Qy 241 TWYQKNDNRNSDETFLRDLPIFFDVLHDPSPFLNYSALSYKWNFGDNTGLFVSNHPTLN 300  
Db 241 TWYQKNDNRNSDETFLRDLPIFFDVLHDPSPFLNYSALSYKWNFGDNTGLFVSNHPTLN 300  
Qy 301 HTYVLNGTFNENLTVOAVPGPCPSPTPS--PSSSTSPSPASSPTLSTPSPSLMPTGY 360  
Db 301 HTYVLNGTFNENLTVOAVPGPCPSPTPS--PSSSTSPSPASSPTLSTPSPSLMPTGY 360  
Qy 361 MELSDISNENCRINRYGYFRATITIVDGIENVIIQVADVPITLQPDNSLMDFIYTC 420  
Db 347 LELSRIPDENCQINRYGHFOATITIVDGIENVIIQVADVPITLQPDNSLMDFIYTC 406  
Qy 421 ATPTEACTIISDPTCOIAONRVCSVAVDELCLLSVRRAPNGSGTYCVNFTLGDDASL 480  
Db 407 SIPTEACTIISDPTCOIAONRVCSVAVDELCLLSVRRAPNGSGTYCVNFTLGDDASL 466  
Qy 481 TSALISIPGKOLGSLRTVNGVLISIGCLAMFVTMTVILLYKKHKYKPIGNCNTRNV 540  
Db 467 TSTLISVDPDRPASPLRMANSLISVGLAIPVTVISLIVYKKHKEYNPISPGNVRS 526  
Qy 541 KGLSVFLSHAKAPFSGRDKDPLLODKP 568  
Db 527 KGLSVFLSHAKAPFSGRDKDPLLODKP 554  
RESULT 4  
AAB11329  
ID AAB11329 standard; Protein; 560 AA.  
XX  
AC AAB11329;  
XX





Query Match		70.7%; Score 2163; DB 23; Length 560;
Best Local Similarity		69.1%; Pred. No. 8 2e-191;
Matches 394; Conservative		75; Mismatches 85; Indels 14; Gaps 1;
QY	1	MESLCGVLVFLLLAAGLPLOAAKFRDVLGHEQYDPHMRNNOLRGWSSDENWDEQLYP 60
DB	1	MECLYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMRHNLQNGSSDENWNEKLYP 60
QY	61	VMRGEGRWKDSWEGGRVQAAALTSDPALVGSNITFVNLVFPKQKEDANGNIYERNK 120
DB	61	VWKRGMKNSWKGGRVQAVLTSDPALVGSNITFAVNLIFPRCKEDANGNIYERNK 120
QY	121	RSDELEASDPYVYNTTGADEWEDNTSGOHLRFDPDGKPPRPHGRKKNVYVFTL 180
DB	121	RNEAGLSADPYVYNTTAWSEDSGNGTQSHHNVFPDGKPPRPHGRKKNVYVFTL 180
QY	181	GOYFQKLGOCARSINTVNLTVGPQVMEVIFRRHGRAYIPISKVDVYVITDQIPFV 240
DB	181	GOYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQVKDVVYVITDQIPFV 240
QY	241	THYCKNDRNSDETFLRDLPIFFDVLHDPSHFLNYSALSKKNFGDNTGLFVSNHTLN 300
DB	241	THYCKNDRNSDETFLRDLPIFFDVLHDPSHFLNYSALSKKNFGDNTGLFVSNHTLN 300
RESULT 6		
AAU83612		
ID		AAU83612 standard; Protein; 572 AA.
AC		AAU83612;
XX		
DT		08-MAY-2002 (first entry)
DE		Human PRO protein, Seq ID No 42.
KW		Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
KW		breast cancer; prostate tumour; rectal tumour; liver tumour;
KW		pericyte cell proliferation; chondrocyte cell proliferation;
KW		tumour necrosis factor-alpha.
OS		Homo sapiens.
XX		
PN		WO200208288-A2.
PD		31-JAN-2002.
XX		
PF		29-JUN-2001; 2001WO-US21066.
XX		
PR		20-JUL-2000; 2000US-219556P.
PR		25-JUL-2000; 2000US-220585P.
PR		25-JUL-2000; 2000US-220605P.
PR		25-JUL-2000; 2000US-220607P.
PR		25-JUL-2000; 2000US-220624P.
PR		25-JUL-2000; 2000US-220638P.
PR		25-JUL-2000; 2000US-220664P.
PR		25-JUL-2000; 2000US-220666P.
PR		25-JUL-2000; 2000US-220668P.
PR		26-JUL-2000; 2000US-220893P.
PR		28-JUL-2000; 2000WO-US20710.
PR		23-AUG-2000; 2000WO-US23522.
PR		24-AUG-2000; 2000WO-US23328.
PR		15-SEP-2000; 2000US-000000P.
PR		10-NOV-2000; 2000WO-US30873.
PR		28-NOV-2000; 2000US-253646P.
PR		01-DEC-2000; 2000WO-US32678.
PR		20-DEC-2000; 2000US-0747259.
PR		20-DEC-2000; 2000WO-US34956.
PR		28-FEB-2001; 2001WO-US06520.
PR		10-MAY-2001; 2001US-0854280.
PR		25-MAY-2001; 2001WO-US17092.
XX		(GETH ) GENENTECH INC.
PA		
XX		
PI		Baker KF, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI		Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX		
XX		WPI; 2002-172001/22.
DR		N-PSDB; ABK33556.
XX		
PT		One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT		useful for treating a PRO related disorder and for diagnosing tumours
PT		such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
PT		tumour or liver tumour -
XX		
XX		Claim 11; Figure 42; 359pp; English.
XX		
CC		The invention relates to one hundred and twenty two nucleic acids
CC		encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
CC		encode human secreted proteins. The PRO nucleic acids, polypeptides,
CC		agonists and antagonists are useful for treating a PRO related disorder.
CC		The PRO polypeptides are useful for diagnosing tumours, especially lung
CC		cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
CC		liver tumour. The PRO polypeptides are useful for stimulating the
CC		proliferation of, or gene expression, in pericyte cells, for stimulating
CC		the proliferation or differentiation of chondrocyte cells, for
CC		stimulating the release of tumour necrosis factor-alpha from human blood,
CC		for stimulating or inhibiting the proliferation of normal human dermal
CC		fibroblast cells. The PRO polypeptide may also be used as molecular
CC		weight markers and for tissue typing. The PRO nucleic acids have
CC		applications in molecular biology, including use as hybridisation probes,
CC		and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO
CC		protein sequences of the invention.
XX		
SQ		Sequence 572 AA;
Query Match		70.6%; Score 2162; DB 23; Length 572;
Best Local Similarity		69.1%; Pred. No. 1e-190;
Matches 396; Conservative		78; Mismatches 87; Indels 12; Gaps 2;
QY	1	MESLCGVLVFLLLAAGLPLOAAKFRDVLGHEQYDPHMRNNOLRGWSSDENWDEQLYP 60
DB	1	MECLYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMRHNLQNGSSDENWNEKLYP 60
QY	61	VMRGEGRWKDSWEGGRVQAAALTSDPALVGSNITFVNLVFPKQKEDANGNIYERNK 120
DB	61	VWKRGMKNSWKGGRVQAVLTSDPALVGSNITFAVNLIFPRCKEDANGNIYERNK 120
QY	121	RSDELEASDPYVYNTTGADEWEDNTSGOHLRFDPDGKPPRPHGRKKNVYVFTL 180
DB	121	RNEAGLSADPYVYNTTAWSEDSGNGTQSHHNVFPDGKPPRPHGRKKNVYVFTL 180
QY	181	GOYFQKLGOCARSINTVNLTVGPQVMEVIFRRHGRAYIPISKVDVYVITDQIPFV 240
DB	181	GOYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQVKDVVYVITDQIPFV 240
QY	241	THYCKNDRNSDETFLRDLPIFFDVLHDPSHFLNYSALSKKNFGDNTGLFVSNHTLN 300
DB	241	THYCKNDRNSDETFLRDLPIFFDVLHDPSHFLNYSALSKKNFGDNTGLFVSNHTLN 300

Db 241 TMFQKNDNRNSDETFKLDLPIMFDVLHDPHSFLNHYSTINYKWSFGDNTGLFVSNHTVYN 300  
QY 301 HTYVLNGTFNMLVOTAVPGCPSPPTSPSSSTSPASSPPTLST-----PSPSLMP 355  
Db 301 HTYVLNGTFNMLTVKAAAPGCPPPPPPP-----PSKPTPSLATTLLKSYDSNTGPG 353  
QY 356 TGYKSMELSDISNENCRINRYGYFRANITIVDGLILEVNIIOVADVPITLQPDNSILMDFI 415  
Db 354 TGNFLELSRPIDENQCINRIYGFQATITIVEGILEVNIIOVADVPITLQPDNSILMDFI 413  
QY 416 VTCKGATPEACTIISDPTCOIAQNRVCSPVADELCLLSVRRAFNGSGTYCVNFTLGGD 475  
Db 414 VTCQGISPTVECTIISDPTCEITQNTVCSPVDVDEMCLLTVRRTEFNGSGTYCVNFTLGGD 473  
QY 476 ASLATSALISTPGKDLGSPRTYNGVLISIGCLAMFVTMTLLYKXKHKTKPIGNCNTR 535  
Db 474 TSLATSLSIVPDRDPASPLRMANSLISVGCLAFVIVISLTVKXKHKTKPIGNCNTR 533  
QY 536 NVYKSGLSVFLSHAKAPFSGDREKDLPLQDK 568  
Db 534 NVYKSGLSVFLNRAKAVFEPGQEKDPLLNQ 566

RESULT 7  
AAW35386  
ID AAW35386 standard; Protein; 376 AA.  
AC AAW35386;  
XX  
XX  
DT 26-FEB-1998 (first entry)  
XX  
DE Murine metastatic nucleic acid sequence product.  
XX  
KW Mouse; murine; tumour; cancer; metastatic sequence; detection;  
KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;  
KW hypertrophy; screening.  
XX  
OS Mus musculus.  
XX  
XX WO9718454-A2.  
XX  
PD 22-MAY-1997.  
XX  
XX 15-NOV-1996; 96WO-US18567.  
XX  
PR 30-JAN-1996; 96US-0594031.  
PR 16-NOV-1995; 95US-0006838.  
XX  
XX (THOM/) THOMPSON T.  
XX  
PI Thompson T;  
XX  
XX WPI; 1997-289397/26..

Identifying tumour metastatic sequences - by introducing transfected cells into host mammal and analysing primary and metastatic sequences by differential display PCR  
PS Disclosure; Fig 12CR; 102pp; English.  
XX  
XX Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mouse embryos. The UGS cells were infected with retroviruses, cultured and implanted under the renal capsule of mice.  
CC Reconstructions were harvested 5 weeks later, when they showed signs of distress from the tumour burden. Metastatised tumours were isolated from a site outside the renal capsule. RNA was isolated from primary tumours and metastases, reverse transcribed and subjected to differential display PCR. The sequences were analysed to obtain metastatic sequences, e.g. the sequence encoding the present sequence. The method can be used to detect, diagnose and treat disorders related to metastasis, or treat malignant or non-malignant disorders, e.g. hyperplasia, dysplasia and hypertrophy. The metastatic sequence can be used to screen a

CC biological sample for metastasis, and it or its expression product  
XX may also be used to treat a metastatic disorder.  
SQ Sequence 376 AA:  
Query Match 21.1%; Score 644.5; DB 18; Length 376;  
Best Local Similarity 38.9%; Pred. No. 1.1e-50;  
Matches 213; Conservative 54; Mismatches 101; Indels 179; Gaps 47;  
QY 16 GLPLQAARFRDVLGHEOYPDHMRNNQLRGWSSDENWEDQLPVWRRGEGRWKDSWEG 75  
Db 5 GAARDAAR-KHDVGRSAY--MRHNN---GWSSD-NDWNY---VWKRGMGRKNSWKG 53  
QY 76 GRVQAALTSDSPALVGSNITFVNLVLEPFCOKEDANGNIVYERNCRSDLELASDPVYNN 135  
Db 54 GRV-AVTSDS--AVGNT--AVN---RC--KANGNY--KNCRN---AGSADVYNN 96  
QY 136 TTGADDEDWEDNTSOGHLRFDPGKPFPRPHGRKKWNPVYVHTLGOYFKLGOCARSVS 195  
Db 97 TAWSG---SDNGTGSNNV-DGR---HHGWRNRYVHTGYK-----GRCSVRVS 139  
QY 196 INTVNLTVGPQVMEVIVFRRHGRAYIPISKYKDVVITDQIPFVTVYKNDNRNSDET 255  
Db 140 VNTAVT-----GMVTVYRRHGRAYV---AVKDVVYTD-----VVTM--KANDRNSDTK- 184  
QY 256 LRDLPIFFDLIHDPSHFLNYSALSYKWNFGDNTGLFVSNHHTLNHYVLTGTFNLT 315  
Db 185 -----DMDVHD-SH--NYST-NYKWS-GDNTG--VSTNHTVNTYV--NGTSN---TV 225  
QY 316 QTAVPGCPSPPTSPSSSTSPASSPPTLSTPSPSLMPTGYKSMELSDISNENCRINR 375  
Db 226 KAA--AGCRS-----KTSAGDNRSDNC--NR 248  
QY 376 YGYFRATITIVDGLILEVNIIOVADVPITLQPDNSLMDFIVTCGATPTACTIISDPTC 435  
Db 249 YGH---ATTG-----VNTDMVWSSD-----VVTG-----GSTVCTSDTC 282  
QY 436 QTAQNRVCSPVADELCLLSVRRAFNGSGTYCVNFTLGGDASLATSALISIPGKDLGSP 495  
Db 283 ---TNTVCS-VDVD---MCTVHRT-NGSGTYCVN--TGDOTSATSTSV-----DRDAS-- 325  
QY 496 LRTVNGVLISIGCLAMFVTMTLLYKXKHKTKPIGNCNTRNVKGLSVFLSHAKAPFS 555  
Db 326 -RMANSA--SVGC-----AVTSVYKHKY-----NNSGNVYRSKG---SVNRAKAV-- 366  
QY 556 RGDREKD 562  
Db 367 -GNKDN 372  
RESULT 8  
AAW35387  
ID AAW35387 standard; Protein; 376 AA.  
XX  
XX AC AAW35387;  
XX  
XX DT 26-FEB-1998 (first entry)  
XX  
XX Murine metastatic nucleic acid sequence product.  
XX  
XX Mouse; murine; tumour; cancer; metastatic sequence; detection;  
XX diagnosis; treatment; metastasis; hyperplasia; dysplasia;  
XX hypertrophy; screening.  
XX  
XX Mus musculus.  
XX WO9718454-A2.  
XX  
PD 22-MAY-1997.  
XX  
XX 15-NOV-1996; 96WO-US18567.  
XX  
XX 30-JAN-1996; 96US-0594031.  
PR



Best Local Similarity 27.8%; Pred. No. 4e-50;  
Matches 176; Conservative 89; Mismatches 191; Indels 178; Gaps 19;

```

QY 54 WDEQLYPVRRGEGRWKDSWEGGRVQAALTSDSALPGNSITFVNILVFPCKQEDANGN 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 43 WNRQLYPEWTEVOG--SNCWRGGQVSLRVINDGPTLIGANASFSIALHFFPGSKVLPDGO 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 IYVERNCRSDLELADSPVYWNWTTGADDEWEDNT---SQ---GQHLR-----PP 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 101 VI-----WANNIINGSOVGGQPVVPQEPDDDACVP 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 158 DGKPPRPGRHKKWNVFVFTLGGYQFKLGOCASRVISINTVNLTVGPQVMEVIVFRRHG 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 DGGPCSPGPKPKRFSVYVWKTGKIQVGLGGPVSRLSIATGHAKLGTHTMEVTVYHRRG 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 218 -RAYTPIKVKVDVYITDQIPIFTVMQKNDNRSSDFTFLRDLPIFDVLIIHDPSPHLY 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 SOSYVPLAHASFTFITDQVPFVSVSQLOALDGETKHLRNHPLIFALQHDHDPGYLEAE 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 277 SAISKWNEFGWTGLFVSNHNTLHNTYVLTNGTFNLTVQTAVP---GPCSP----- 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 253 ADSLYTWDFGDGTGLISRALDVTHTYLESGSVTAQVVLQAAIPLVSCGSSPVPGTTDGY 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 327 -----TPSSSSTSPS-----PASSPSPTLS----- 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 MPTAEAPGTTSQGTTTEVGGTTPGQMPPTQPSGTTVQMPPTTEVTATSEQMLTSAVID 372
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 -----TPSPS-----LMPTGYKSMELSDISNE----- 369
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 373 TTLAEVSTTEGGTTPSPSTTGAQATTTEGPDASPLLLPTOSSTGGISPLLDLDTIMLV 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 370 -----NCRINRYGYERATITIVOGIILEVNIIQVADVPITLQPDNSLMDFTVCKGATP 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 433 KRQVPLDCVLYRYGFSALDIVOGIESAELLQA--VPF---SEGDAFELVSCGGGLP 486
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 424 TRACTIISDPTCQIAQNRVCSPVAVBELCLLSVRRAF--NGSGTYCVNFTLGDDASLALTS 482
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 487 KEACMDISSPGCGPPAQRLCQSPVPPDCQLVLHQVLKGGSGTYCLNVSLADANSIAYAS 546
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 483 ALIISTPGKDLG---SPLRTVNGVLISIGCLAMFVTMTILLYKKKHATYKPIGNCRTNVVK 539
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 547 TQLVVPQDGGGLGQAPL-----LVGILLVAVVVLASLIHRH-----LK 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 540 GKGLSVFLSHAKAPF-----SRGDREKDPPLL 565
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 587 KQGSYSQMPHGSHWLRPLPVFVRARGLGENSPLL 620
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
ID AAR78646
XX AAR78646 standard; Protein; 561 AA.
XX AAR78646;
XX
XX
XX 22-JAN-1996 (first entry)
XX Melanoma associated antigen gp100.
XX
XX Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
XX Identification; tumour; gp100.
XX
XX Homo sapiens.
XX
XX EP668350-A1.
XX
XX 23-AUG-1995.
XX
XX 14-FEB-1995; 95EP-0200348.
XX
XX 21-DEC-1994; 94EP-0203709.
XX 16-FEB-1994; 94EP-0200337.
XX
XX (ALKU ) AKZO NOBEL NV.

```

XX AC AAY31577;  
 XX DT 21-DEC-1999 (first entry)  
 XX DE Human melanoma antigen gp100.  
 XX DE Melanoma antigen gp100; human; antigen presentation;  
 KW dendritic cell; adoptive immunotherapy; cancer; therapy; vaccine.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Peptide 154..162  
 FT /note= "minimal essential epitope"  
 FT Peptide 209..217  
 FT /note= "minimal essential epitope"  
 FT Peptide 280..288  
 FT /note= "minimal essential epitope"  
 FT Peptide 457..466  
 FT /note= "minimal essential epitope"  
 FT Peptide 476..485  
 FT /note= "minimal essential epitope"  
 XX WO9947102-A2.  
 XX DT 23-SEP-1999.  
 XX PF 19-MAR-1999; 99WO-US06031.  
 XX PR 20-MAR-1998; 98US-0078880.  
 XX PA (GENZ ) GENZYME CORP.  
 XX PI Nicolette CA, Kaplan J;  
 XX DR WPI; 1999-590956/50.  
 XX PT Preparing cells for use as cancer vaccines and in adoptive  
 XX immunotherapy -  
 XX PS Disclosure; Page 44-46; 55pp; English.  
 XX CC The present sequence represents the human melanoma antigen gp100,  
 CC a melanocyte differentiation antigen specifically recognised by  
 CC HLA-A2 restricted tumour-infiltrating lymphocytes of melanoma  
 CC patients. The invention provides methods for immunotherapy, in  
 CC particular for inducing an immune response against an antigen in a  
 CC patient. Genetically modified antigen-presenting cells (APC) that  
 CC are more potent presenters of exogenous peptide than the parental  
 CC antigen-presenting cells are used. These APCs lack an effective  
 CC endogenous TAP (transporter associated with antigen processing)  
 CC activity and present exogenous antigen on the major  
 CC histocompatibility complex class I (MHC-I) molecule. Suitable  
 CC exogenous antigens include a tumour antigen, such as a minimal  
 CC essential epitope of gp100, which can complex with MHC-I. The  
 CC genetically modified APCs are useful for inducing an immune  
 CC response (claimed) against an antigen in a patient (adoptive  
 CC immunotherapy), especially as vaccines against cancer in mammals,  
 CC preferably humans. The cells are also useful for expanding  
 CC populations of immune effector cells, preferably cytotoxic T  
 CC lymphocyte cells.  
 XX SQ Sequence 661 AA;  
 Query Match 20.5%; Score 627.5; DB 20; Length 661;  
 Best Local Similarity 27.1%; Pred. No. 9.5e-49;  
 Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;  
 QY VLVFLLAAGLPLQAAK--RFRDVLGHEQYPDHMRNNLCRGSSDENEDQLPYVWRR 64  
 DB 9 LLHLAVIGALLAVGATKVPKQNDLQ-----VSRLR-----TKANRQLYPEWT- 53

QY 65 GEGRWKOSGEGRYQAALTSDSPALVGSNITFVNVLFPRCKEDANGNIVYERN-CRSD 123  
 DB 54 -EAQLDCWRGQVSLKVSNDGPTLIGANASFSIALNFGSKVLDPDQGVVWNTIING 112  
 QY 124 LELASDPVYVYNTTGADEDEDWNTSOGHLRFPDQKPPRHRKKKNNFVVFHTLQGY 183  
 DB 113 SOVMGGQVYPOET--DDA-----CIFPDGQPCPSGWSQKRSFVYVWKTWQGY 159  
 QY 184 FQKLGQCSARYSINTVNLTVGPQVMEVIFRRHG-RAYIPISKVKDVYVITQIPIFVTM 242  
 DB 160 WQVLGGPVSGLSIGTGRAMLGTHMEVTVYHRRGSYSYVPLAHSSSAFTITQVFPESVY 219  
 QY 243 YOKNDRNSSDETFRLDLPFVDVLIHDPHFLNYSISYKWNFGDNTGLFVNNHTLNT 302  
 DB 220 SOLRALDGGNKHFLRNOPLTFALQLHDPSPGYLAEDLSYTWDFGSSGTLISRALVVTHT 279  
 QY 303 VYVNGTFENLTVOTAVP-----GPCP----- 324  
 DB 280 YLEPGPVTAQVVLQAALPLISCGSSPVPGTIDGHRPTAEAPNTAGQVTTVEVGTTPGQ 339  
 QY 325 SPTSPSSST-----SPSPASSPSTLSTPSPSLMP-----TG 357  
 DB 340 APTAEPSTTSVQVPTTEVISTAPVQMPTAESTGMTPEKVPVSEVMGTTLABKSTPEATG 399  
 QY 358 YKSMELSDI-----SNE----- 369  
 DB 400 WTPAEVSIVLGGTTAAQVTTTEWETARELPIPEPEGPDASSIMSTESITGSLGPLLD 459  
 QY 370 -----NCRINRYGYPRATITVDGILEVNIITQVADVPITLQPDNSLMDPI 415  
 DB 460 GTATLRLVKRQVPLDCVLYRYGSPSVTLDIVQGIESAEILQA---VPSGEGD--AFELT 513  
 QY 416 VTCKGATPTEACTIISOPTCOIAQNRVCSPVAVDELCLLSVRRAF-NGSGTYCVNFTLGD 474  
 DB 514 VSCQGLPKFAKMEISSPCQPPAQRLCQVPLSPACQILVHLKGGSGTYCLNVSLAD 573  
 QY 475 DASIALTSALISIPKQDILG-SPLRTVNGVVISIGCLAMFVTWYTIILYKK 523  
 DB 574 TNSLAVVSTQIMPGQENGALQGVPLIVGIL-----LVLMVAVLASLIYRR 618  
 RESULT 12  
 AAB47540  
 ID AAB47540 standard; Protein; 661 AA.  
 XX AAB47540;  
 DT 13-DEC-2001 (first entry)  
 DE Human melanoma antigen gp100.  
 KW Major histocompatibility complex; MHC; human; melanoma antigen; gp100;  
 KW HLA-A2 binding domain; mutation; antigen presenting cell; vaccine;  
 KW immune effector cell; cancer; antibody.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Binding-site 209..217  
 FT /label= HLA-A2 binding domain  
 XX WO200170767-A2.  
 XX DT 27-SEP-2001.  
 XX PF 19-MAR-2001; 2001WO-US08919.  
 XX PR 20-MAR-2000; 2000US-190750P.  
 XX PR 12-DEC-2000; 2000US-255019P.  
 XX PA (GENZ ) GENZYME CORP.  
 XX PI Nicolette CA;



Db 220 SQLRALDGGNKHFLRNQPLTFALQHDPSGYLAEDLSYTWDFGDSSTGLSRLVVTHT 279  
 QY 303 YVLNGTFNLTQVAVP-----GPCP----- 324  
 Db 280 YLEPGPVTAQVVLQAAIPLTSCGSSPVPGTTCGHRTAAPNTAGQVPTTEVGTTPGQ 339  
 QY 325 SPTPSSST-----SPSPASSPPTLSTPSPSLMP-----TG 357  
 Db 340 APTAEPSTTSVQVPTTEVISTAPVQMPAESTGMTPEKVPVSEVGMGTTLAEMSTPEATG 399  
 QY 358 YKSMELSDI-----SNE----- 369  
 Db 400 MTPAEVSIWLSGTTAAQVTTTETWETTARELPPEPEGPDASSIMSTESITGLSLPLD 459  
 QY 370 -----NCRINRYGYFRATITIVDGIENVIIQVADVPIPTLPQDNLMDFI 415  
 Db 450 GTATLRLVKROVPLDCVLYRGSFVTLDIVOGIESAEILQA-----VPSGEGD--AFELT 513  
 QY 416 VTCKGATPTEACTIISDPTCQIAQNRVCSVAVDELCLLSVRRAF-NGSGTYCVNFTLGD 474  
 Db 514 VSCGGGLPKACMEISSPGCPQPAQLCQVLPSPACQLVHLQILKGGSGTYCLNVSLAD 573  
 QY 475 DASIALTSALISIPKDLG-SPLRTVNGVLISIGCLAMFVTWVTLIYKK 523  
 Db 574 TNSLAVVSTQLIMPQGEAGLQGVPLIVGIL-----LVLMMAVVLASLIYRR 618

RESULT 14  
 AAU84803  
 ID AAU84803 standard; Protein: 661 AA.  
 XX AAU84803;  
 AC  
 DT 08-MAY-2002 (first entry)  
 XX Human gp100 consensus sequence.  
 DE  
 KW Saviene; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200190197-A1.  
 XX  
 XX 29-NOV-2001.  
 XX  
 XX 25-MAY-2001; 2001WO-AU00622.  
 XX  
 XX 26-MAY-2000; 2000AU-0007761.  
 XX  
 XX (AUSU ) UNIV AUSTRALIAN NAT.  
 XX  
 XX Thomson SA, Ranshaw IA;  
 XX  
 XX WPI; 2002-147575/19.  
 XX  
 PT New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response  
 PT against a pathogen or cancer  
 XX  
 PS Example 3; Fig 27; 364pp; English.  
 XX  
 CC The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for

designing the synthetic polypeptides. The synthetic polypeptides and  
 polynucleotides are referred to as a saviene. The synthetic polypeptide is  
 useful for modulating immune responses preferably directed against a  
 pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,  
 CC Salmonella, streptococcus, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence is a consensus sequence for a parent protein used to design a  
 CC saviene of the invention.  
 XX  
 SQ Sequence 661 AA;  
 Query Match 20.5%; Score 627.5; DB 23; Length 661;  
 Best Local Similarity 27.1%; Pred. No. 9.5e-49;  
 Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;  
 QY 7 VLVFLILAAAGLPLQAAK--RFRDVLGHEQYDPDHMRNNQLRGNSSDENDEQLYVWRR 64  
 Db 9 LLHLAVIGALLAVGATKVPKQMDLW-----VSRQLR-----TKANRQLYPEWT- 53  
 QY 65 GEGRWKDSWEGGRVQAALTSALVGSNTITFVNLVFPFCQKEDANGNTIVYERN-CRSD 123  
 Db 54 -EAQRLDCMRGQVSLKVSNDGPTLIGANASFSIALNFGSGVLPDQGVVNNYNTIING 112  
 QY 124 LELASDPVYVNWTTGADDEDWEDNTSQGHLPFGDPGPPRPHGRKKNNEVYVHTLQGY 183  
 Db 113 SQVWGQPVYPOET--DDA-----CIFPDGPGCPSGNSQKRSFVYVKTWQGY 159  
 QY 184 FOKLGCSARVSNITVNLTVGPQWNEVIVFRRHG-RAYIPISKVKQVYVTTDIPFTVM 242  
 Db 160 WQVLGPGVSLGIGTGRAMLGTHMEVTVYHRRGSRVYPLAHSSSAFTTDOQPFESV 219  
 QY 243 YKNDNRNSDETFRLDLPIDFFDLIHDPSHFLNYSALSKWNFCNDTGLFVSNHNTLHT 302  
 Db 220 SQLRALDGGNKHFLRNQPLTFALQHDPSGYLAEDLSYTWDFGDSSTGLSRLVVTHT 279  
 QY 303 YVLNGTFNLTQVAVP-----GPCP----- 324  
 Db 280 YLEPGPVTAQVVLQAAIPLTSCGSSPVPGTTCGHRPTAEAPNTAGQVPTTEVGTTPGQ 339  
 QY 325 SPTPSSST-----SPSPASSPPTLSTPSPSLMP-----TG 357  
 Db 340 APTAEPSTTSVQVPTTEVISTAPVQMPAESTGMTPEKVPVSEVGMGTTLAEMSTPEATG 399  
 QY 358 YKSMELSDI-----SNE----- 369  
 Db 400 MTPAEVSIWLSGTTAAQVTTTETWETTARELPPEPEGPDASSIMSTESITGLSLPLD 459  
 QY 370 -----NCRINRYGYFRATITIVDGIENVIIQVADVPIPTLPQDNLMDFI 415  
 Db 460 GTATLRLVKROVPLDCVLYRGSFVTLDIVOGIESAEILQA-----VPSGEGD--AFELT 513  
 QY 416 VTCKGATPTEACTIISDPTCQIAQNRVCSVAVDELCLLSVRRAF-NGSGTYCVNFTLGD 474  
 Db 514 VSCGGGLPKACMEISSPGCPQPAQLCQVLPSPACQLVHLQILKGGSGTYCLNVSLAD 573  
 QY 475 DASIALTSALISIPKDLG-SPLRTVNGVLISIGCLAMFVTWVTLIYKK 523  
 Db 574 TNSLAVVSTQLIMPQGEAGLQGVPLIVGIL-----LVLMMAVVLASLIYRR 618  
 RESULT 15  
 AAU09695  
 ID AAU09695 standard; Protein: 661 AA.  
 XX AAU09695;  
 AC

XX DT 26-MAR-2002 (first entry)  
 XX DE Human melanoma antigen gp100.  
 XX KW Human: anti-melanoma compound; melanoma antigen gp100; APC; MHC;  
 KW immune effector cell; antigen presenting cell; anti-cancer;  
 KW major histocompatibility complex; gp100 tumour antigen; cytostatic.  
 XX OS Homo sapiens.  
 XX PN WO200192294-A2.  
 XX PD 06-DEC-2001.  
 XX PF 21-MAY-2001; 2001WO-US16417.  
 XX PR 31-MAY-2000; 2000US-208955P.  
 XX PR 09-FEB-2001; 2001US-267877P.  
 XX PA (GENZ ) GENZYME CORP.  
 XX PI Nicolette CA;  
 XX DR WPI; 2002-106301/14.  
 XX DR N-PSDB; AAS14396.  
 XX PT Novel anti-melanoma compound or peptide useful for inducing immune  
 PT response in a subject, for treating melanoma, as components of  
 PT anti-cancer vaccines and to expand immune effector cells specific for  
 PT cancers -  
 XX PS Claim 11; Page 65-66; 69pp; English.  
 XX CC The present invention relates to anti-melanoma compounds comprising a  
 CC peptide sequence based on human melanoma antigen gp100. Also described  
 CC are antibodies that recognise and bind to these compounds.  
 CC polynucleotides that encode these compounds, and immune effector cells  
 CC that have been raised in vitro or in vivo in the presence of an antigen  
 CC presenting cell (APC) that presents the compound. Such an APC may be  
 CC the major histocompatibility complex (MHC) molecule. The anti-melanoma  
 CC compounds are useful for inducing an immune response in a subject, by  
 CC delivering the compound to the subject in the context of an MHC molecule  
 CC which presents the compound on the surface of an APC. The anti-melanoma  
 CC compound is delivered as a polynucleotide that encodes it. The compounds  
 CC are useful to generate antibodies that specifically recognise and bind  
 CC to them, for the treatment of melanoma, as components of anti-cancer  
 CC vaccines, and to expand immune effector cells that are specific for  
 CC cancers characterised by expression of gp100 tumour antigen, melanoma.  
 CC The compounds are also useful in diagnostic methods for such diseases.  
 CC The present sequence represents human melanoma antigen gp100.  
 XX SQ Sequence 661 AA;  
 Query Match 20.5%; Score 627.5; DB 23; Length 661;  
 Best Local Similarity 27.1%; Pred. No. 9.5e-49;  
 Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;  
 QY 7 VLVFLLLAAGLPLOAAK--RFRDVLGHEQYPDHMRNNOLRGWSSDENEDWSQLYPMWR 64  
 DB 9 LHLAVIGALLAVGATKVPKQDWLG-----VSRQLR-----TKAWNRLYPEWT- 53  
 QY 65 GEGRWKDSWEGRRVOAALSDSPALVGSNTFFVNLVFFPRCKEDANGNIVYERN-CRSD 123  
 DB 54 -EAQLDCWRGGQVSLKVSNDGPTLIGANASFIALNFPGSQKVLPGDGVWNNNTIING 112  
 QY 124 LELASDPVYVNTTCADEDWEDNTSGOHLRFPDGKPPRPHGRKKKNFVYVHTLGOY 183  
 DB 113 SQWGGQVPYQET--DDA-----CIFDGGPCPGSGWSQKRSFYVWKWTGOY 159  
 QY 184 FQKLGCCSARYSINTVNLVGPQVMEIVPFRHG-RAYIPISKVKDVIYITDQIPFVTM 242  
 DB 160 MQVLGGPVSGLSIGTGRAMLGTHTMEVTVVHRGSRSYVPLAHSSSAFTITDQVPFVS 219

QY 243 YQKDRNSDSETFRLDLPIFFDVLTHDPSHEFNYSATSYKWNFGDNTGLFVNNHNLNHT 302  
 DB 220 SQLRALDGGNKHFLRNQPLTALQHLDPSPGYLAELADLSYTWDFGSSGTLISRALVVTHT 279  
 QY 303 YVLNGTFENLTVQTAVP-----GPCP----- 324  
 DB 280 YLEPGVTVQVVLQAAIPLTSCGSSPVPVGTDDGHRPTAEAPNTTAGOVPTTEVVGTPGQ 339  
 QY 325 SPTPSPSSST-----SPSPASSPSTLSTPSPSLMP-----TG 357  
 DB 340 APTAEPSTGTSVQVPTTEVISTAPVQMPATBSTGMTPEKVPVSEVMGTTLAEMSTPEATG 399  
 QY 358 YKSMELSDI-----SNE----- 369  
 DB 400 MTPAEVSIIVVLSGTTAAQVTTTEWVETTARELPPEPEGPPASSIMTESITGSLGPLLD 459  
 QY 370 -----NCRINRYGYFRATITIVDGIILEVNIQVADVPIPTLPDQNSLMDFI 415  
 DB 460 GTATRLVKRQVPLDCVLYRYGFSVTLDIYQGIESAELQA-----VPSGEGD--APELT 513  
 QY 416 VTCKGATPTEACTIISDPTCOIAQNRVCSVAVDELCLLSVRRAF-NGSGTYCVNFTLGD 474  
 DB 514 VSCQGLPKREACWEISSPGCQPPAQRICQVLPSPACQLVLHQILKGGSGTYCLNWSLAD 573  
 QY 475 DASLALTSALISIPGKDLG-SPLRTVNGVGLISIGCLAMFVTMTVLTILYKK 523  
 DB 574 TNSLAVVSTQLIMPQGEAGLGQVPLIVGIL-----LVLMAVVVLASLIYRR 618

Search completed: April 7, 2003, 16:09:59  
 Job time : 81 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 16:08:12 ; Search time 28 Seconds  
(without alignments)  
601.068 Million cell updates/sec

Title: US-09-943-075A-2  
Perfect score: 3061  
Sequence: 1 MESICGVLLVFLLAAGLPLQ.....PFSRGDRKDPDLQDKPWL 572

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/PCTUS-COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2700	88.2	574	4	US-09-383-586-36
2	2163	70.7	560	1	Sequence 36, Appl
3	2163	70.7	560	1	Sequence 90, Appl
4	644.5	21.1	376	1	Sequence 225, App
5	644.5	21.1	376	1	Sequence 100, App
6	627.5	20.5	661	2	Sequence 102, App
7	627.5	20.5	661	4	Sequence 121, App
8	626.5	20.5	661	2	Sequence 27, Appl
9	626.5	20.5	661	2	Sequence 27, Appl
10	626.5	20.5	661	2	Sequence 27, Appl
11	626.5	20.5	661	2	Sequence 27, Appl
12	602	19.7	668	1	Sequence 6, Appl
13	157	5.1	71	1	Sequence 98, Appl
14	131.5	4.3	521	1	Sequence 3, Appl
15	123	4.0	1426	4	Sequence 43, Appl
16	118.5	3.9	616	4	Sequence 47, Appl
17	116.5	3.8	4302	4	Sequence 8, Appl
18	116.5	3.8	4339	4	Sequence 6, Appl
19	113	3.7	2972	4	Sequence 2, Appl
20	113	3.7	3118	4	Sequence 1, Appl
21	112.5	3.7	4302	3	Sequence 5, Appl
22	112.5	3.7	4303	2	Sequence 2, Appl
23	110	3.6	438	2	Sequence 59, Appl
24	108.5	3.5	1751	4	Sequence 44, Appl
25	107	3.5	490	4	Sequence 2, Appl
26	107	3.5	536	4	Sequence 21, Appl
27	107	3.5	555	4	Sequence 15, Appl

28	107	3.5	555	4	US-09-292-225-18
29	103.5	3.4	610	1	US-07-821-717B-6
30	103.5	3.4	610	1	US-08-119-262B-6
31	103.5	3.4	610	1	US-08-135-929A-11
32	103.5	3.4	610	1	US-08-234-265A-11
33	100.5	3.3	2071	4	US-09-415-522-6
34	99.5	3.3	496	2	US-08-224-482-2
35	99.5	3.3	533	1	US-08-040-548-1
36	99.5	3.3	533	1	US-08-466-344-1
37	99.5	3.3	533	6	5206152-2
38	95.5	3.1	542	4	US-09-189-462-6
39	95.5	3.1	766	3	US-08-539-205A-4
40	95.5	3.1	2319	1	US-08-212-133A-8
41	95.5	3.1	2319	1	US-08-474-503-6
42	95.5	3.1	2319	2	US-08-670-707A-6
43	95.5	3.1	2319	4	US-09-037-601-6
44	95.5	3.1	2319	4	US-09-315-179-6
45	95.5	3.1	2319	4	US-09-523-656-28

ALIGNMENTS

RESULT 1

US-09-383-586-36  
; Sequence 36, Application US/09383586  
; Patent No. 6242419  
; GENERAL INFORMATION:  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Abernethy, Nevin  
; APPLICANT: Onrust, Rene  
; APPLICANT: Kumble, Anand  
; APPLICANT: Murison, Greg  
; TITLE OF INVENTION: Compounds isolated from stromal cells  
; TITLE OF INVENTION: and methods for their use  
; FILE REFERENCE: 11000.1037c1  
; CURRENT APPLICATION NUMBER: US/09/383,586  
; CURRENT FILING DATE: 1999-08-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 35  
; LENGTH: 574  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-383-586-36

Query Match	88.2%	Score	2700;	DB 4;	Length	574;
Best Local Similarity	88.8%	Pred. No.	6.9e-255;			
Matches	507;	Conservative	20;	Mismatches	42;	Indels
2;	Gaps	1;				
QY	1	MESICGVLLVFLLAAGLPLQAAKRFDRVLGHEQYPDHMRNNQLRGWSSDENEDQLYP	60			
DB	1	MESICGVLLVFLLAAGLPLQAAKRFDRVLGHEQYPDHMRNNQLRGWSSDENEDQLYP	60			
QY	61	VWRGGRKWSWEGRGVQAALTSDSPALVGSNTTFVYVNLVFPKCKEDANGNIVYERNK	120			
DB	61	VWRGGRKWSWEGRGVQAALTSDSPALVGSNTTFVYVNLVFPKCKEDANGNIVYERNK	120			
QY	121	RSDLKASDPYVYNTWTGADDEDWEDNTSQOHLRFPDGFPPHGRKKNVYVYFHTL	180			
DB	121	RNDGLTSLHYVYNTWTGADDEDWEDNTSQOHLRFPDGFPPHGRKKNVYVYFHTL	180			
QY	181	GOYFQKLGCCSARVSNVTNLTGVPQVMEVAFVRRHGRAYIPISKVDYVITDQIPFV	240			
DB	181	GOYFQKLGCCSARVSNVTNLTGVPQVMEVAFVRRHGRAYIPISKVDYVITDQIPFV	240			
QY	241	TMCKNDNRSSDETFRLDLPFFOVLIIHDPGHFLNYSALSYKWNFGDNTGLFVSNHNTLN	300			
DB	241	TMCKNDNRSSDETFRLDLPFFOVLIIHDPGHFLNYSALSYKWNFGDNTGLFVSNHNTLN	300			
QY	301	HTVYLGCTFNFLVQTAVPGCPSPPTSS--PSSSTSPSPASSPTLSTPSPSLMPTGY	358			
DB	301	HTVYLGCTFNFLVQTAVPGCPSPPTSS--PSSSTSPSPASSPTLSTPSPSLMPTGY	358			



Db 1 MECLYFELGFLAARLPLDAARFHDVLGNRPSAYMRHNLNGWSSDENDAKELYP 60  
Qy 61 VWRGEGRWKDSWEGRVQAALTSDSPALVGSNITFVNLVPPROCKEDANGNIVYERNK 120  
Db 61 VWRGDMRWKNSWGGVQAVLTSDSPALVGSNITFVNLVPPROCKEDANGNIVYERNK 120  
Qy 121 RSDLELASDPYVYNTTGADEWEDNTSGOHLRFPPDGKPPRPHGRKKNFVVFHTL 180  
Db 121 RNEAGLSADPYVYNTTAWSESDSGENGTSQSHHNVFPDGKPPRPHGRWNNFYVFTL 180  
Qy 181 GOYFQKLGOCARSIVNTVLTGPOVMEVIFERRHGRAYIPISKVDYVITDQIPIFV 240  
Db 181 GOYFQKLGOCARSIVNTVLTGPOVMEVIFERRHGRAYIPISKVDYVITDQIPIFV 240  
Qy 241 TMYOKNDRNSSDETFLRDLPIFFDLVLIHDPSPFLNYSATSKWNGFDMGTGLFVSNHHTLN 300  
Db 241 TMYOKNDRNSSDETFLRDLPIFFDLVLIHDPSPFLNYSATSKWNGFDMGTGLFVSNHHTLN 300  
Qy 301 HTYVLTGTFNLTQVAVPCPSPTPSPSSSTSPSPASSPTLSTPSPSLMPTGYKS 360  
Db 301 HTYVLTGTFNLTQVAVPCPSPTPSPSSSTSPSPASSPTLSTPSPSLMPTGYKS 360  
Qy 361 MELSISNCRINRYGFRATITVDGILEVNIIOVADVPITLQPDNSLMDFTVTCG 420  
Db 361 MELSISNCRINRYGFRATITVDGILEVNIIOVADVPITLQPDNSLMDFTVTCG 420  
Qy 421 ATPTEACTIISDPTCQIAQNRVCSVAVDELCLLSVRRAFNGSGTYCNFTLGGDASLAL 480  
Db 421 ATPTEACTIISDPTCQIAQNRVCSVAVDELCLLSVRRAFNGSGTYCNFTLGGDASLAL 480  
Qy 481 TSALISIPGKDLASFTVNGVLISIGCLAMFVTMTILLYKKHKTYPIGNCTRNIVYK 540  
Db 481 TSALISIPGKDLASFTVNGVLISIGCLAMFVTMTILLYKKHKTYPIGNCTRNIVYK 540  
Qy 541 KGLSVFLSHAKAPSRDREKDLQDK 568  
Db 541 KGLSVFLSHAKAPSRDREKDLQDK 568  
Qy 568 KGLSVFLSHAKAPSRDREKDLQDK 568  
Db 568 KGLSVFLSHAKAPSRDREKDLQDK 568

RESULT 4  
US-08-594-031-100  
; Sequence 100, Application US/08594031  
; Patent No. 5783182  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/594,031  
; FILING DATE: 30-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: 0A146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700

; TELEFAX: 202-639-7890  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 100:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 376 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; US-08-594-031-100  
  
Query Match 21.1%; Score 644.5; DB 1; Length 376;  
Best Local Similarity 38.9%; Pred. No. 1.9e-54;  
Matches 213; Conservative 54; Mismatches 101; Indels 179; Gaps 47;  
  
Qy 16 GLPLQAARFRDVLGHEOYPDHMRNOLRGWSSDENDEQLYPVWRRGEGRWKDSWEG 75  
Db 5 GAARDAAR-HDVGNRSAI---MRHNN---GWSSD-NDWNY---VMKRGDMRKNSWKG 53  
Qy 76 GRVQAALTSDSPALVGSNITFVNLVPPROCKEDANGNIVYERNCRSDLEASDPYVYNN 135  
Db 54 GRV--AVTSDS--AVGNT--AVN---RC--KDANGNY--KNCRN--AGSADIYVYNN 96  
Qy 136 TFGADEDEWEDNTSGOHLRFPPDGKPPRPHGRKKNFVVFHTLGOYFQKLGOCARSIV 195  
Db 97 TAWSD---SDNGTSGSHHNV-DCK---HHGMRNRNIVYHTGYK-----GRCSVRVS 139  
Qy 196 INTVNLTVGPOVMEVIFERRHGRAYIPISKVDYVITDQIPIFVMTYOKNDRNSSDET 255  
Db 140 VNTAVNT-----GMVTVYRHRGRAYV---AVKDYVYVD---VYTM--KANDRSSDTK- 184  
Qy 256 LRDLPIFFDLVLIHDPSPFLNYSATSKWNGFDMGTGLFVSNHHTLNHYVLTGTFNLT 315  
Db 185 -----DMDVHD-SH--NYST-NYKWS-GDNTG--VSTNHTVNTYV--NGTSN---TV 225  
Qy 316 QTAVPGPCPSPTPSPSSSTSPSPASSPTLSTPSPSLMPTGYKSMELSDISNCRINR 375  
Db 226 KAA--AGCRS-----KTSAGDNRSDNC--NR 248  
Qy 376 YGYFRATITVDGILEVNIIOVADVPITLQPDNSLMDFTVTCGATPTACTIISDPTC 435  
Db 249 YGH-----ATTVG-----VNMTDVMWSSD-----VVTC-----GSTVCTSDTC 282  
Qy 436 QIAQNRVCSVAVDELCLLSVRRAFNGSGTYCNFTLGGDASLALTSIPGKDLGSP 495  
Db 283 ---TWTVCV--VDVD---MCTVRRT--NGSGTYCVN--TGDTSATSTSV-----DRDAS-- 325  
Qy 496 LRTVNGVLISIGCLAMFVTMTILLYKKHKTYPIGNCTRNIVYKGLSVFLSHAKAPFS 555  
Db 326 -RWANSA--SVGC-----AVTVSVYKHKY-----NNSGNVYRSKG---SVNRKAV-- 366  
Qy 556 RGDREK 562  
Db 367 -GRDKN 372  
  
RESULT 5  
US-08-594-031-102  
; Sequence 102, Application US/08594031  
; Patent No. 5783182  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC



```
QY 243 YKNDNRSSDETFLRDLPIFFDVLLHDPKSHFLNYSALSYKKNFGDNTGLFVSNHNTLHNT 302
D 220 SOLRALDGGNKHFLRNQPLTALQHDPSGYLAELADLSYTWDFGSSGTLISRALVYVHT 279
QY 303 YVLNGTFNFLTVOAVP-----GPCP----- 324
D 280 YLEPGPVTAQVLAQAIPLTSCGSSVPVGTGDHRTAEAPNTTAGOVPTTEVVGTPGQ 339
QY 325 SPTPSPSSST-----SPSPASSPSTLTSPSPSLMP-----TG 357
D 340 APTAEPSTTSVQVPTTEVISTAPVQMPAETGTPKVPVSEVVGTTLAEMSTPEATG 399
QY 358 YKSMELSDI-----SNE----- 369
D 400 MTPAEVSIVVLSTGTTAAQVTTETVETARELPPEPEGPDASSIMSTESITGSLGPLLD 459
QY 370 -----NCRINRYGYFRATITVDGILEVNIQVADVPITPLQPDNSLMDFI 415
D 460 GTATLRLVKRQVPLDCVLYRYGFSVTLDIVOGIESAEILQA-----VPSGEGD--AFELT 513
QY 416 VTCKGATPEACTIISDPTCQIAONRVCSPVAVDELCLLSVRRAF--NGSGTYCVNFTLGD 474
D 514 VSCGGLPKKACHEISPGCQPPAQRLCQVLPSPACQLVHLHILKGGSGTYCLNVSLAD 573
QY 475 DASLALTSALISIPGKDLG--SPLRTVNGVLISIGCLAMFVMTVITLLYKK 523
D 574 TNSLAVVSTQIMPGOEAGLGQVPLIVGIL-----LVLMAVVLASLIYRR 618
```

## RESULT 7

US-09-267-439-121

Sequence 121, Application us/09267439

Patent No. 6270778

GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,

APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND

TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 126

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN &amp; FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/267,439

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/417,174

FILING DATE: 05-APR-1995

APPLICATION NUMBER: US/08/231,565

FILING DATE: 22-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: CAROL M. GRUPPI

REGISTRATION NUMBER: 37,341

REFERENCE/DOCKET NUMBER: 2026-4124US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 121:

SEQUENCE CHARACTERISTICS:

LENGTH: 661

TYPE: amino acid

```
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Protein
US-09-267-439-121
```

## Query Match

20.5%; Score 627.5; DB 4; Length 661;

Best Local Similarity 27.1%; Pred. No. 2.2e-52;

Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

```
QY 7 VLVFLLAAGLPLQAAK--RFRDVLGHEQYPDHMRNNOLRGKSSDENEDQLYPMVRR 64
D 9 LLHLAVIGALLAVGATKVPNRQDWLG-----VSRQLR-----TKANNQLYPEWT- 53
QY 65 GGRKWKDSWEGRGVRAALTSDSPALVGSNITFVNVLVFPKCKEDANGNIVERN--CRSD 123
D 54 -BAQLDLCVRGGSVLKVSNDGPTLLIGANASFSIALNFFGSKQVLPDGVIVNNTIING 112
QY 124 LELASDPYVYNNVTGADDEDWEDNTSOGHLRFPDCKPPEPRPHGRKKNFVVEHTLGOY 183
D 113 SQVMGGQPVYPOET--DDA-----CIFPDGPGCPGSGWSOKRSFVYVWKWGOY 159
QY 184 FOKLGCOSARVINTVNLTVGPMVEVIVFRHGG--RAYIPISKVDVYVITDQIFIFVTM 242
D 160 WQVJGGPVSGLSIGTGRAMLGTHTMEVTVYHRRGSRVYVPLAHSSSAFTITDQVPFSVSV 219
QY 243 YKNDNRSSDETFLRDLPIFFDVLLHDPKSHFLNYSALSYKKNFGDNTGLFVSNHNTLHNT 302
D 220 SOLRALDGGNKHFLRNQPLTALQHDPSGYLAELADLSYTWDFGSSGTLISRALVYVHT 279
QY 303 YVLNGTFNFLTVOAVP-----GPCP----- 324
D 280 YLEPGPVTAQVLAQAIPLTSCGSSVPVGTGDHRTAEAPNTTAGOVPTTEVVGTPGQ 339
QY 325 SPTPSPSSST-----SPSPASSPSTLTSPSPSLMP-----TG 357
D 340 APTAEPSTTSVQVPTTEVISTAPVQMPAETGTPKVPVSEVVGTTLAEMSTPEATG 399
QY 358 YKSMELSDI-----SNE----- 369
D 400 MTPAEVSIVVLSTGTTAAQVTTETVETARELPPEPEGPDASSIMSTESITGSLGPLLD 459
QY 370 -----NCRINRYGYFRATITVDGILEVNIQVADVPITPLQPDNSLMDFI 415
D 460 GTATLRLVKRQVPLDCVLYRYGFSVTLDIVOGIESAEILQA-----VPSGEGD--AFELT 513
QY 416 VTCKGATPEACTIISDPTCQIAONRVCSPVAVDELCLLSVRRAF--NGSGTYCVNFTLGD 474
D 514 VSCGGLPKKACHEISPGCQPPAQRLCQVLPSPACQLVHLHILKGGSGTYCLNVSLAD 573
QY 475 DASLALTSALISIPGKDLG--SPLRTVNGVLISIGCLAMFVMTVITLLYKK 523
D 574 TNSLAVVSTQIMPGOEAGLGQVPLIVGIL-----LVLMAVVLASLIYRR 618
```

## RESULT 8

US-08-417-174-27

Sequence 27, Application US/08417174

Patent No. 5844075

GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,

APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND

TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 126

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN &amp; FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-08-417-174-27

Query Match 20.5%; Score 626.5; DB 2; Length 661;  
Best Local Similarity 27.1%; Pred. No. 2.8e-52;  
Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

QY 7 VLVLALLAGLPQAAK--RFRDVLGHEQYPDHMRNNQLRGWSSDENWDEQLYPVWRR 64  
DB 9 LLHLAVIGALLAVGATKVPKQDNLG-----VSRLR-----TKAQRQLYPEWT- 53

QY 65 GEGRWKDSWEGGRVQAALTSDSPALVGSNTTFVNNLVFPRCKEDANGNIVYERN-CRSD 123  
DB 54 -EAQRUDCWRRGGVSLKVSNDGPTLLIGANASFSIALNFPQSKVLPDGOVIVNNITING 112

QY 124 LEASDPYVYNNMTTGADEWDENTSGQHLRFPDCKPPRPHGRKKWNFVYVHTLGOY 183  
DB 113 SQWGGQPVYPOET--DDA-----CIPDGGPCPSGWSQKRSFYVWKTGWQY 159

QY 184 FQKLGQCSARVSINTNLTVGQVMEIVFRRHG-RAYIPISKVKDYYVITQIPFVTM 242  
DB 160 WQPLGGPVSGLSIGTGRAMLGTHTMEVYHRRGRSRYVPLAHSSSAFTITDQVPFSVS 219

QY 243 YQKNDNRSSDETFLRDLPIFDVLHDPHFNLNYSALSYKWNFGDNTGLFVSNHHTLHIT 302  
DB 220 SQLRALDGGNKHFLRNQPLTFALQHDPSGYLAEADLSYTWDFGSSGTLISRALVYVHT 279

QY 303 YVLNGTFNFLTQVATVP-----GPCP----- 324

DB 280 YLEPGPTAQVVLQAAIPLSCGSSPVGTYDGRHPTAEAPNTAGQVPTTEVGTTPGQ 339

QY 325 SPTSPSSST-----SPSPASSPSTLTSTPSSLMP-----TG 357  
DB 340 APTAEPSTTSVQVPTTEVISTAPVQMTAESTGMTPEKVPFVSEVMGTTLAEMSTPEATG 399

QY 358 YKSMELSDI-----SNE----- 369

DB 400 MTPAEVSVLWLSGTTAAQVTTETWETARELPPEPEGPDASSIMTESITGSLGLPLD 459

QY 370 -----NCRINRYGYFRATITIVDGIILEYNIIVQADVPIPTLOPNSLMDFI 415  
DB 460 GTATLRLVKRQVPLDCVLYRGSFVTLDIQVIGESAEILQA-----VPSGEGD--APELT 513

QY 416 VTCGATPTEACTIISDPTQAIQNRVCSVPVAVDELCLLSVRRAF-NGSGTCYCNFTLGD 474  
DB 514 VSCGGGLPKEACMEISSPGOPPAQRILCQVPLPSACOLVHLQTLKSGSGTYCLNLSLAD 573

QY 475 DASLALTSALISIPGKDLG-SPLRTVNGVLTSIGCLAMFVTMTVITLLYKK 523

DB 574 TNSLAVVSTQIMPGQEAQLGQVPLIVGIL-----LVLMAVVLASLIYRR 618

RESULT 9  
US-08-231-565A-27  
Sequence 27, Application US/08231565A  
Patent No. 5874560  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565A  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-08-231-565A-27

Query Match 20.5%; Score 626.5; DB 2; Length 661;  
Best Local Similarity 27.1%; Pred. No. 2.8e-52;  
Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

QY 7 VLVLALLAGLPQAAK--RFRDVLGHEQYPDHMRNNQLRGWSSDENWDEQLYPVWRR 64  
DB 9 LLHLAVIGALLAVGATKVPKQDNLG-----VSRLR-----TKAQRQLYPEWT- 53

QY 65 GEGRWKDSWEGGRVQAALTSDSPALVGSNTTFVNNLVFPRCKEDANGNIVYERN-CRSD 123  
DB 54 -EAQRUDCWRRGGVSLKVSNDGPTLLIGANASFSIALNFPQSKVLPDGOVIVNNITING 112

QY 124 LEASDPYVYNNMTTGADEWDENTSGQHLRFPDCKPPRPHGRKKWNFVYVHTLGOY 183  
DB 113 SQWGGQPVYPOET--DDA-----CIPDGGPCPSGWSQKRSFYVWKTGWQY 159

QY 184 FQKLGQCSARVSINTNLTVGQVMEIVFRRHG-RAYIPISKVKDYYVITQIPFVTM 242  
DB 160 WQPLGGPVSGLSIGTGRAMLGTHTMEVYHRRGRSRYVPLAHSSSAFTITDQVPFSVS 219

QY 243 YQKNDNRSSDETFLRDLPIFDVLHDPHFNLNYSALSYKWNFGDNTGLFVSNHHTLHIT 302  
DB 220 SQLRALDGGNKHFLRNQPLTFALQHDPSGYLAEADLSYTWDFGSSGTLISRALVYVHT 279

QY 303 YVLNGTFNFLTQVATVP-----GPCP----- 324

```
Db 280 VLEPGPVTAQVVVLAQAAIPLTSCGSSVPVGGTTDGHRTAEAPNTTAGQVPTTEVVGTTTGG 339
QY 325 SPTSPSSST-----SPSPASSPSTLTSTPSPSLMP-----TG 357
Db 340 APTAEPSTTSVQVPTTEVISTAPVQMPAESTGMTPEKVPVSEVNGTTLAEMSTPEATG 399
QY 358 YKSMELSDI-----SNE-----369
Db 400 MTPAEVSIIVLGGTTAAQVTTTETARELPIPEPGDASSIMSTESITGSLGLD 459
QY 370 -----NCRINRYGYFRATITIVDGIILEVNIIOVADVPITLQPDNSLMDFI 415
Db 460 GTATLRLVKRQVPLDCVLYRYGSFVTLDIVOGIESAEILQA-----VPSGEGD--AFELT 513
QY 416 VTCKGATPEACTIISDPTQIAQNRVCSVAVDELCLLSVRRAF--NGSGTYCVNFTLGD 474
Db 514 VSCQGLPKACMEISSPGCOPPAORLCQVPLSPACQLVHLQILKGGSGTYCLNVSLAD 573
QY 475 DASLALTSALISPGKDLG--SPLRTYNGVLISIGCLAMFVTWVTTILYKK 523
Db 574 TNSLAVSTQLIMPQEGAGLQGVPLIVGIL-----LVLMAVVLASLIYRR 618
```

## RESULT 10

```
US-09-007-961-27
; Sequence 27, Application US/09007961
; Patent No. 5994523
; GENERAL INFORMATION:
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
; APPLICANT: STEVEN A.
; TITLE OF INVENTION: MELANOMA ANTIGENS AND
; THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
; METHODS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007,961
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/231,565
; FILING DATE: 22-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 661
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Protein
US-09-007-961-27
Query Match 20.5%; Score 626.5; DB 2; Length 661;
Best Local Similarity 27.1%; Pred. No. 2.8e-52;
```

```
Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;
QY 7 VLVFLLAAGLPLOAK--RFRDVLGHEQYDPDHRENNQLRGHSSDENEDWDBOLIPVRRR 64
Db 9 LUHLAVIGALLAVGATKVPKRNODWLG-----VSQRJR-----TKAWNQLYPEMT- 53
QY 65 GECRWKDSWEGRGVQAALTSQSPALVGSNITFVVNLVFPKQKEDANGINIVYERN-CRSD 123
Db 54 -ENQRLDWCWGGVSLKVSNDGPTLIGANASFSIALNPGSKVLPDQGVVWNNITING 112
QY 124 LEASDPVYVWNTTCADDEDWEDNTSOGHLRFPDGPFRPHGRKKWKNVYVFTLGOY 183
Db 113 SQVWGGQVYPOET--DDA-----CFFDGGPCPSGWSOKRSFVYVWKTWGOY 159
QY 184 FQLQGCARSIVNTVLTGPOVMEVIFVRRHG--RAYIPISKVKDVIYITDQIPIFTVM 242
Db 160 WQFLGGPVSGLSIGTGRAMLGTHTMEVTYVHRRGSRVYVPLAHSSSAFTITDQVPFSV 219
QY 243 YQKNDNSDDETFRLDLPFFDVLHDPHPLNYSAISYKKNFCDNTGLFVSNHHTLHT 302
Db 220 SQLRALDGGKHFLRNQPLTFALQHDPSGYLAEDLSYTWDFGSSGTLISRALVVTHT 279
QY 303 YVINGTFENLTQVATVP-----GPCP-----SNE-----369
Db 280 YLEPGPVTAQVVVLAQAAIPLTSCGSSVPVGGTTDGHRTAEAPNTTAGQVPTTEVVGTTTGG 339
QY 325 SPTSPSSST-----SPSPASSPSTLTSTPSPSLMP-----TG 357
Db 340 APTAEPSTTSVQVPTTEVISTAPVQMPAESTGMTPEKVPVSEVNGTTLAEMSTPEATG 399
QY 358 YKSMELSDI-----SNE-----369
Db 400 MTPAEVSIIVLGGTTAAQVTTTETARELPIPEPGDASSIMSTESITGSLGLD 459
QY 370 -----NCRINRYGYFRATITIVDGIILEVNIIOVADVPITLQPDNSLMDFI 415
Db 460 GTATLRLVKRQVPLDCVLYRYGSFVTLDIVOGIESAEILQA-----VPSGEGD--AFELT 513
QY 416 VTCKGATPEACTIISDPTQIAQNRVCSVAVDELCLLSVRRAF--NGSGTYCVNFTLGD 474
Db 514 VSCQGLPKACMEISSPGCOPPAORLCQVPLSPACQLVHLQILKGGSGTYCLNVSLAD 573
QY 475 DASLALTSALISPGKDLG--SPLRTYNGVLISIGCLAMFVTWVTTILYKK 523
Db 574 TNSLAVSTQLIMPQEGAGLQGVPLIVGIL-----LVLMAVVLASLIYRR 618
RESULT 11
US-09-267-439-27
; Sequence 27, Application US/09267439
; Patent No. 6270778
; GENERAL INFORMATION:
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
; APPLICANT: STEVEN A.
; TITLE OF INVENTION: MELANOMA ANTIGENS AND
; THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
; METHODS
; NUMBER OF SEQUENCES: 126
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/267,439
; FILING DATE:
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,174
; FILING DATE: 05-APR-1995
; APPLICATION NUMBER: US/08/231,565
; FILING DATE: 22-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4124US1
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 661
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Protein
; US-09-267-439-27

Query Match 20.5%; Score 626.5; DB 4; Length 661;
Best Local Similarity 27.1%; Pred. No. 2.8e-52;
Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

Qy 7 VLVELLLAAGLPLOAAK--RFRDVLGHEQYPDHMRNNQLRGWSSDENEDQLYPVWRR 64
Db 9 LLHLAVIGALLAVGATKVPKRNQDWLG-----VSRQLR-----TKAWNRQLYPEWT- 53

Qy 65 GEGRWKSDWEGKRVQAALTSPPALVGSNTITFVNVLPFPCQKEDANGNTVYERN-CRSD 123
Db 54 -EAQLDRCWGGVSLKVSNDGPTLIGANASFALNFPQSQKVLDPQGVIVWNNIING 112

Qy 124 LELASDPVYVNTTGADDEDWEDTSCQHLRFDPGKPEPRPHGRKKNVYVVFHTLGOY 183
Db 113 SQWVGQPVYPOET--DDA-----CIPDGGCPGSGWSQKRSFYVVKWTGQI 159

Qy 184 FOKLGCASARVINTVNLTVGPQWMEVIVFRRHG-RAYIPISKVKDYVITDQIPFVFM 242
Db 160 WQFLGPGVSGLSIGTGRAMLGTHMTVTYVYHRRGSRVYVPLAHSSAFTITDQVPFSVS 219

Qy 243 YQKNDNRNSDFTFLDLPTFFDVLTHDPSSHFLNYSATSYKWNFGDNTGLPVSNNHFLNHT 302
Db 220 SQLRDLGGKHELRNQLPTALQLHDFSGYLAEDLSYTWDFGDSSTGLISRALVYVHT 279

Qy 303 YVUNGTFNENLVOTAVP-----GPCB-----GCPB----- 324
Db 280 YLEFGPVTAQVVLQAAIPLTSCGSSPVPGTTDGHRTAEAPNTTAGOVPTTEVVGTPGQ 339

Qy 325 SPTPPSSST-----SPSPASSPPTLSTPSPSLMP-----TG 357
Db 340 APTAEPFGTISVQVPTTEVISTAPVOMTAESTGMTPEKVPVSEVMGTTLAENSTPEATG 399

Qy 358 YKSMELSDI-----SNE----- 369
Db 400 MTPAEYSIVVLGGTAAQVTTTEWVETARELPPEPEGPDPASSIMSTESITGSLGPLLD 459

Qy 370 -----NCRINRGYFATITVDGILEVNIIOVADVPITPTLPDPSNLMDFI 415
Db 460 GTATLRLVKRQVPLDCLVLYRGFSFVTLDIVOGIESAEILQA-----VPSGEGD--AFELT 513

Qy 416 VTCKGATPTEACTIISDPTCCQAQRNRCVSPVAVDELCLLSVRRAF-NGSGTYCVNFTLGD 474
Db 514 VSCGGPLPKEACMEISSPCCPPAQRQLCQVLPSPACQLVLHQLKGGSTYCLNLSLAD 573

Qy 475 DASLATSAISIPKDKIG-SPLRTVNGVLISIGCLAMFTWMTILYKK 523
Db 574 TNSLAYVSTQLIMPQGEAGLGQVPLVINGIL-----LVLMAVVLASLIYR 618

RESULT 12
US-07-891-9426-6

```

```

; Sequence 6, Application US/07891942G
; Patent No. 5679511
; GENERAL INFORMATION:
; APPLICANT: Kwon, Byoung Se
; TITLE OF INVENTION: CDNA CLONES FOR HUMAN TYROSINASE AND FOR
; TITLE OF INVENTION: A REGULATORY PROTEIN IN THE MELANIN PROTEIN PATHWAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christopher A. Michaels, Barnard, Brown &
; ADDRESSEE: Michaels
; STREET: 306 East State Street; Suite 220
; CITY: Ithaca
; STATE: NY
; COUNTRY: USA
; ZIP: 14850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,942G
; FILING DATE: 01-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/915,753
; FILING DATE: 06-OCT-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/362,847
; FILING DATE: 07-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaels, Christopher A
; REGISTRATION NUMBER: 34,390
; REFERENCE/DOCKET NUMBER: IND1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 607-273-1711
; TELEFAX: 607-273-2609
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PUBLICATION INFORMATION:
; AUTHORS: Kwon, Byoung Se
; AUTHORS: Chintamani, Chaya
; AUTHORS: Kozak, Christine A
; AUTHORS: Copeland, Neal G
; AUTHORS: Gilbert, Debra J
; AUTHORS: Jenkins, Nancy
; AUTHORS: Barton, David
; AUTHORS: Francke, Uta
; AUTHORS: Kobayashi, Yvonne
; AUTHORS: Kim, Kack K
; TITLE: A melanocyte-specific gene, Pmel 17, maps
; TITLE: near the silver coat color locus on mouse
; TITLE: chromosome 10 and is in a syntenic region on human
; TITLE: chromosome 12
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 88
; PAGES: 9228-9232
; DATE: October-1991
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 668
; US-07-891-942G-6

```

```

Query Match 19.7%; Score 602; DB 1; Length 668;
Best Local Similarity 26.8%; Pred. No. 7e-50;
Matches 175; Conservative 95; Mismatches 212; Indels 170; Gaps 19;

Qy 7 VLVELLLAAGLPLOAAK--RFRDVLGHEQYPDHMRNNQLRGWSSDENEDQLYPVWRR 64
Db 9 LLHLAVIGALLAVGATKVPKRNQDWLG-----VSRQLR-----TKAWNRQLYPEWT- 53

```





US-08-276-213-3

Query Match 4.3%; Score 131.5; DB 1; Length 521;  
Best Local Similarity 20.9%; Pred. No. 0.00042;  
Matches 63; Conservative 51; Mismatches 84; Indels 103; Gaps 13;  
QY 230 YVITDOI-PIFVTYOKNDNRSSDETFURDLPIFFDVLHDPHFNLSAISKWNF--- 285  
DB 269 YLFNOMIAFVWVGEFTTQSTDTQTLWKTL-----VQYLRPT--AQYGADSFQWTFWSW 321  
QY 286 ----GNTGLFVSNHHTLNHYTVLNGTFNENLTVOTAV-----PGPCSPSP 328  
DB 322 NPDSGDTGILKDDQWTD--TVKDG---YLAPIKSSIFDPVGCASAPSPSPSPSP 376  
QY 329 SPSSSTSPSPASSPPTLSTSPSLMPTGYKSMELSDISNENCRINRYGYFRATITVDG 388  
DB 377 SPSPASRPTPTPTPTAS--PTPLTPTA----- 404  
QY 389 ILEVNIQVADVPITLQPDNSLMDFICTCKGATPTACTIISDPTCQIAONRVCSPVAV 448  
DB 405 -----TPTPT-----ASPTSPSTAAGACTASYQ-----VNS 432  
QY 449 DELCLLSVRANFGSG-----TYCVNFTLGDASLALT-SALISIPGKDLGSLPLTVNGV 502  
DB 433 DMGNGFTTVAVTNSGVSATKTWTVSWTFEGNQITITNSWNAAVTQNGOSVTARNMSYNNV 492  
QY 503 L 503  
DB 493 I 493

RESULT 15

US-09-136-574A-43  
: Sequence 43, Application US/09136574A  
: Patent No. 6294366  
: GENERAL INFORMATION:  
: APPLICANT: Farrington, Graham K.  
: Anderson, Paige  
: Gibbs, Moreland  
: Bergquist, Peter  
: Daniels, Roy  
: Morgan, Hugh W.  
: Williams, Diane P.  
: TITLE OF INVENTION: Compositions and Methods for  
: Treating Cellulose Containing  
: Cellulase Enzyme Compositions  
: NUMBER OF SEQUENCES: 49  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Howson and Howson  
: STREET: Spring House Corporate Center, P.O. Box 457  
: CITY: Spring House  
: STATE: PA  
: COUNTRY: USA  
: ZIP: 19477  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/136,574A  
: FILING DATE: 19-Aug-1998  
: CLASSIFICATION: <Unknown>  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/932,571  
: FILING DATE: September 19, 1997  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Bak, Mary E.  
: REGISTRATION NUMBER: 31,215  
: REFERENCE/DOCKET NUMBER: 1997US001/CIP  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 215-540-9200  
: TELEFAX: 215-540-5818  
: TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1426 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: NO. 6294366e  
SEQUENCE DESCRIPTION: SEQ ID NO: 43:  
US-09-136-574A-43  
Query Match 4.0%; Score 123; DB 4; Length 1426;  
Best Local Similarity 25.6%; Pred. No. 0.015;  
Matches 44; Conservative 27; Mismatches 31; Indels 70; Gaps 10;  
QY 315 VOTAVPGPCSP-----TPSPSSSTSPSPASSPPTLS---TPSPSLMPTGYKSMEL 363  
DB 952 IPTTTPTPTPTVTPTSTPTSPVSSSTPTPTATPTPTPTPTPTPTPTPTPTPTPTPT 1009  
QY 364 SDISNEN-----CRINRYGYFRATITVDGILEVN---IIQVAD- 399  
DB 1010 TDDTNDWLFAQGNKIVDKGKPVWLTVGNWFGFNTGT-NVFDGWSNLSKALAEIANR 1068  
QY 400 -----VPI-----PTLQPDNSL--MDFIV-TCK 419  
DB 1069 GFNLLRVPISAEILLNWSKGIYKPKPNINYVNPPEGLTSLVDFVFKTK 1120  
Search completed: April 7, 2003, 16:12:44  
Job time : 34 secs



```
Db 181 GQYFQKLGCCSARVSINTVNLTVGPQVMEVIVFRRHGRAYIPIISKVKDYYVITDQIPIFV 240
Qy 241 TMYQKNDNRNSDETFRLDLPIDFVLIHDPHFNLNYSALSYKKNFNGDNTGLFVSNHHTLN 300
Db 241 TMYQKNDNRNSDETFRLDLPIDFVLIHDPHFNLNYSALSYKKNFNGDNTGLFVSNHHTLN 300
Qy 301 HTVVLNGTFNLTQTAVTAVGPGCPSPSPSSSTSPSPASSPSTLSTPSPSLMPTGYKS 360
Db 301 HTVVLNGTFNLTQTAVTAVGPGCPSPSPSSSTSPSPASSPSTLSTPSPSLMPTGYKS 360
Qy 361 MELSDISNENCRINRYGYFRATITIVDGIILEVNIIOVADVPIPTLOPDSNLMDFIVTCG 420
Db 361 MELSDISNENCRINRYGYFRATITIVDGIILEVNIIOVADVPIPTLOPDSNLMDFIVTCG 420
Qy 421 ATPTACTIISDPTCOIAQNRVCSVAVDELCLLSVRRAFNGSGTYCVNFTLGDGASLAL 480
Db 421 ATPTACTIISDPTCOIAQNRVCSVAVDELCLLSVRRAFNGSGTYCVNFTLGDGASLAL 480
Qy 481 TSALISIPGKDLGSLRTVNGVLIISIGCLAMPVMTVITILLYKKHKTYKPIGNCNTRNVYK 540
Db 481 TSALISIPGKDLGSLRTVNGVLIISIGCLAMPVMTVITVITILLYKKHKTYKPIGNCNTRNVYK 540
Qy 541 KGLSVFLSHAKAPFSRGDRKDPQLQKPWML 572
Db 541 KGLSVFLSHAKAPFSRGDRKDPQLQKPWML 572

RESULT 2
US-09-943-075A-5
; Sequence 5, Application US/09943075A
; Patent No. US20020151486A1
; GENERAL INFORMATION:
; APPLICANT: Popoff, Steven N.
; APPLICANT: Safado, Favez F.
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: Osteoactivin Protein and Nucleic Acids Encoding the Same,
; FILE REFERENCE: 71369.262
; CURRENT APPLICATION NUMBER: US/09/943.075A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 60/229,006
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Mouse
US-09-943-075A-5
```

```
Query Match 88.4%; Score 2706; DB 10; Length 574;
Best Local Similarity 89.0%; Pred. No. 1.2e-215;
Matches 508; Conservative 20; Mismatches 41; Indels 2; Gaps 1;

Qy 1 MESLCGVLFLLAAGLPLOAAKFRDVLGHEQYPDHMRNQLRGWSSDENEDQLYP 60
Db 1 MESLCGVLFLLAAGLPLOAAKFRDVLGHEQYPDHMRNQLRGWSSDENEDQLYP 60
Qy 61 VWRGEGRWKDSWEGGRVQAALTSDSPALVGSNITFVNLVFPFCOKEDANGNIYVERNC 120
Db 61 VWRGEGRWKDSWEGGRVQAALTSDSPALVGSNITFVNLVFPFCOKEDANGNIYVERNC 120
Qy 121 RSDLELASDPYVYVNTWTGADDEDWEDNTSQOHLRFDPDGKPPRPHGKKWKFVYVPHTL 180
Db 121 RNDLGLTSDLVHYVNTWTAGDDGWDGTSRSHQLRFPDRRPFPRPHGKKWKFVYVPHTL 180
Qy 181 GQYFQKLGCCSARVSINTVNLTVGPQVMEVIVFRRHGRAYIPIISKVKDYYVITDQIPIFV 240
Db 181 GQYFQKLGCCSARVSINTVNLTVGPQVMEVIVFRRHGRAYIPIISKVKDYYVITDQIPIFV 240
Qy 241 TMYQKNDNRNSDETFRLDLPIDFVLIHDPHFNLNYSALSYKKNFNGDNTGLFVSNHHTLN 300
Db 241 TMYQKNDNRNSDETFRLDLPIDFVLIHDPHFNLNYSALSYKKNFNGDNTGLFVSNHHTLN 300
```

```
Db 241 TMSQKNDNRNSDETFRLDLPIDFVLIHDPHFNLNYSALSYKKNFNGDNTGLFVSNHHTLN 300
Qy 301 HTVVLNGTFNLTQTAVTAVGPGCPSPSPSSSTSPSPASSPSTLSTPSPSLMPTGY 358
Db 301 HTVVLNGTFNLTQTAVTAVGPGCPSPSPSSSTSPSPASSPSTLSTPSPSLMPTGY 360
Qy 359 KSMELSDISNENCRINRYGYFRATITIVDGIILEVNIIOVADVPIPTLOPDSNLMDFIVTC 418
Db 361 KSMELSDISNENCRINRYGYFRATITIVDGIILEVNIIOVADVPIPTLOPDSNLMDFIVTC 420
Qy 419 KGATPTACTIISDPTCOIAQNRVCSVAVDELCLLSVRRAFNGSGTYCVNFTLGDGASL 478
Db 421 KGATPTACTIISDPTCOIAQNRVCSVAVDELCLLSVRRAFNGSGTYCVNFTLGDGASL 480
Qy 479 ALTSALISIPGKDLGSLRTVNGVLIISIGCLAMPVMTVITILLYKKHKTYKPIGNCNTRNV 538
Db 481 ALTSTLIISIPGKDPDSPLRAVNGVLIISIGCLAMPVMTVITILLYKKHKAYKPIGNCNTRNV 540
Qy 539 KGKGLSVFLSHAKAPFSRGDRKDPQLQDKP 569
Db 541 KGKGLSVLLSHAKAPFSRGDRKDPQLQDKP 571

RESULT 3
US-09-823-038A-36
; Sequence 36, Application US/09823038A
; Patent No. US20020058335A1
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compositions Isolated From Stromal Cells
; FILE REFERENCE: 11000.1037c3
; CURRENT APPLICATION NUMBER: US/09/823.038A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Mouse
US-09-823-038A-36

Query Match 88.2%; Score 2700; DB 10; Length 574;
Best Local Similarity 88.8%; Pred. No. 3.8e-215;
Matches 507; Conservative 20; Mismatches 42; Indels 2; Gaps 1;

Qy 1 MESLCGVLFLLAAGLPLOAAKFRDVLGHEQYPDHMRNQLRGWSSDENEDQLYP 60
Db 1 MESLCGVLFLLAAGLPLOAAKFRDVLGHEQYPDHMRNQLRGWSSDENEDQLYP 60
Qy 61 VWRGEGRWKDSWEGGRVQAALTSDSPALVGSNITFVNLVFPFCOKEDANGNIYVERNC 120
Db 61 VWRGEGRWKDSWEGGRVQAALTSDSPALVGSNITFVNLVFPFCOKEDANGNIYVERNC 120
Qy 121 RSDLELASDPYVYVNTWTGADDEDWEDNTSQOHLRFDPDGKPPRPHGKKWKFVYVPHTL 180
Db 121 RNDLGLTSDLVHYVNTWTAGDDGWDGTSRSHQLRFPDRRPFPRPHGKKWKFVYVPHTL 180
Qy 181 GQYFQKLGCCSARVSINTVNLTVGPQVMEVIVFRRHGRAYIPIISKVKDYYVITDQIPIFV 240
Db 181 GQYFQKLGCCSARVSINTVNLTVGPQVMEVIVFRRHGRAYIPIISKVKDYYVITDQIPIFV 240
Qy 241 TMYQKNDNRNSDETFRLDLPIDFVLIHDPHFNLNYSALSYKKNFNGDNTGLFVSNHHTLN 300
Db 241 TMSQKNDNRNSDETFRLDLPIDFVLIHDPHFNLNYSALSYKKNFNGDNTGLFVSNHHTLN 300
Qy 301 HTVVLNGTFNLTQTAVTAVGPGCPSPSPSSSTSPSPASSPSTLSTPSPSLMPTGY 358
Db 301 HTVVLNGTFNLTQTAVTAVGPGCPSPSPSSSTSPSPASSPSTLSTPSPSLMPTGY 358
```

Db 301 HTYVLNCTFNINLTAVTAVPCPPSPSTPPPPPLTLSTPSPSLMPTGY 360  
 QY 359 KSMELSDISNENCRINRYGYFRATITVDGILEVNIQVADVPITLQPDNSLMDFIYTC 418  
 Db 361 KSMELSDISNENCRINRYGYFRATITVEGILEVNIQVADVPITLQPDNSLMDFIYTC 420  
 QY 419 KGATPTEACTIISDPTCOIAONRVCSFVAVDELCLLSVRRAFNGSGTYCVNFTLGDDASL 478  
 Db 421 KGATPMEACTIISDPTCOIAONRVCSFVAVDGLCLLSVRRAFNGSGTYCVNFTLGDDASL 480  
 QY 479 ALTSALISIPCKDLSPLRVNGVLISIGCLAMFVTWVITLLYKKHKKYKPIGNCNRNV 538  
 Db 481 ALTSALISIPCKDLSPLRVNGVLISIGCLAMFVTWVITLLYKKHKKYKPIGNCNRNV 540  
 QY 539 KGKGLSVFLSHAKAPFSGRGREKDPDLQDKP 569  
 Db 541 KGKGLSVLLSHAKAPFSGRGREKDPDLQDKP 571

RESULT 4  
 US-10-039-272-2  
 ; Sequence 2, Application US/10039272  
 ; Patent No. US20020168653A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RAMESHWAR, Pranela  
 ; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE  
 ; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey  
 ; CURRENT APPLICATION NUMBER: US/10/039,272  
 ; CURRENT FILING DATE: 2001-10-20  
 ; PRIOR APPLICATION NUMBER: US 60/241,881  
 ; PRIOR FILING DATE: 2000-10-20  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 560  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-039-272-2

Query Match 70.7%; Score 2163; DB 9; Length 560;  
 Best Local Similarity 69.4%; Pred. No. 9.8e-171;  
 Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

QY 1 MESICGVLVFLLLAAGLPQAARFRDVLGHEQYDPHMRNQLRGWSSDENEDQLYP 60  
 Db 1 MECLYFGLFLLAARLPDAARFHDVLGNRPSTAYMREHNQNLGWSDDENEDQLYP 60  
 QY 61 VMRGEGRWKDSWEGGRVQAALTSDSPALVGSNITFVNLVFPKQKEDANGNIYERNK 120  
 Db 61 VMRGDMRWKNSWKGGRVQAALTSDSPALVGSNITFVNLVFPKQKEDANGNIYERNK 120  
 QY 121 RSDLELASDPYVYNTWTGADDEWEDNTSQOHLRFDPDGKPPRPHGRKKWNVYVFTL 180  
 Db 121 RNEAGLSADPYVYNTWTGADDEWEDNTSQOHLRFDPDGKPPRPHGRKKWNVYVFTL 180  
 QY 181 GYFQKLGRCVSRVSVANTVNLGQVMEVYVRRHGRAYIPTSKVKDYVITDQIPFV 240  
 Db 181 GYFQKLGRCVSRVSVANTVNLGQVMEVYVRRHGRAYIPTSKVKDYVITDQIPFV 240  
 QY 241 TMQKNDNRSSDDETLFDLPIMFDVLTHDPSHFLNYSATSKWNGDNGLFVSNHNTLN 300  
 Db 241 TMQKNDNRSSDDETLFDLPIMFDVLTHDPSHFLNYSATSKWNGDNGLFVSNHNTLN 300  
 QY 301 HTYVLNCTFNINLTAVTAVPCPPSPSTPPSPSSSTSPSPSSSTSPSPSSSTSPSPSS 360  
 Db 301 HTYVLNCTFNINLTAVTAVPCPPSPSTPPSPSSSTSPSPSSSTSPSPSSSTSPSPSS 360  
 QY 361 MELSDISNENCRINRYGYFRATITVDGILEVNIQVADVPITLQPDNSLMDFIYTC 420  
 Db 361 MELSDISNENCRINRYGYFRATITVDGILEVNIQVADVPITLQPDNSLMDFIYTC 420  
 QY 421 ATPTTEACTIISDPTCOIAONRVCSFVAVDELCLLSVRRAFNGSGTYCVNFTLGDDASL 480  
 Db 421 ATPTTEACTIISDPTCOIAONRVCSFVAVDELCLLSVRRAFNGSGTYCVNFTLGDDASL 480

Db 407 SIFTEVCTIISDPTCEITQNTVCSFVDVDEMCLLTVRRTEFNGSGTYCVNLTGDDTSLAL 466  
 QY 481 TSALLSIPCKDLSPLRVNGVLISIGCLAMFVTWVITLLYKKHKKYKPIGNCNRNVK 540  
 Db 467 TSLLSVDRDPAFPLMANALISVGCCLAFVTVISLLVYKKHKEYNPIENSPGNVRS 526  
 QY 541 KGLSVFLSHAKAPFSGRGREKDPDLQDK 568  
 Db 527 KGLSVFLNRAKAVFPFGNQEKDPLLNQ 554

RESULT 5  
 US-09-735-705-225  
 ; Sequence 225, Application US/09735705  
 ; Patent No. US20020052329A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Hoeken, Nancy  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Fanger, Neil  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; FILE REFERENCE: 210121.455C14  
 ; CURRENT APPLICATION NUMBER: US/09/735,705  
 ; CURRENT FILING DATE: 2000-12-12  
 ; NUMBER OF SEQ ID NOS: 419  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 225  
 ; LENGTH: 550  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-735-705-225

Query Match 70.7%; Score 2163; DB 10; Length 560;  
 Best Local Similarity 69.4%; Pred. No. 9.8e-171;  
 Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

QY 1 MESICGVLVFLLLAAGLPQAARFRDVLGHEQYDPHMRNQLRGWSSDENEDQLYP 60  
 Db 1 MECLYFGLFLLAARLPDAARFHDVLGNRPSTAYMREHNQNLGWSDDENEDQLYP 60  
 QY 61 VMRGEGRWKDSWEGGRVQAALTSDSPALVGSNITFVNLVFPKQKEDANGNIYERNK 120  
 Db 61 VMRGDMRWKNSWKGGRVQAALTSDSPALVGSNITFVNLVFPKQKEDANGNIYERNK 120  
 QY 121 RSDLELASDPYVYNTWTGADDEWEDNTSQOHLRFDPDGKPPRPHGRKKWNVYVFTL 180  
 Db 121 RNEAGLSADPYVYNTWTGADDEWEDNTSQOHLRFDPDGKPPRPHGRKKWNVYVFTL 180  
 QY 181 GYFQKLGRCVSRVSVANTVNLGQVMEVYVRRHGRAYIPTSKVKDYVITDQIPFV 240  
 Db 181 GYFQKLGRCVSRVSVANTVNLGQVMEVYVRRHGRAYIPTSKVKDYVITDQIPFV 240  
 QY 241 TMQKNDNRSSDDETLFDLPIMFDVLTHDPSHFLNYSATSKWNGDNGLFVSNHNTLN 300  
 Db 241 TMQKNDNRSSDDETLFDLPIMFDVLTHDPSHFLNYSATSKWNGDNGLFVSNHNTLN 300  
 QY 301 HTYVLNCTFNINLTAVTAVPCPPSPSTPPSPSSSTSPSPSSSTSPSPSSSTSPSPSS 360  
 Db 301 HTYVLNCTFNINLTAVTAVPCPPSPSTPPSPSSSTSPSPSSSTSPSPSSSTSPSPSS 360  
 QY 361 MELSDISNENCRINRYGYFRATITVDGILEVNIQVADVPITLQPDNSLMDFIYTC 420  
 Db 361 MELSDISNENCRINRYGYFRATITVDGILEVNIQVADVPITLQPDNSLMDFIYTC 420  
 QY 421 ATPTTEACTIISDPTCEITQNTVCSFVDVDEMCLLTVRRTEFNGSGTYCVNLTGDDTSLAL 466  
 Db 421 ATPTTEACTIISDPTCEITQNTVCSFVDVDEMCLLTVRRTEFNGSGTYCVNLTGDDTSLAL 466

QY 421 ATPTACTIISDPTCQIAONRVCSFVAVDELCLLSVRRAPNGSGTYCVNFTLGGDASLAL 480  
Db 407 SIPTVECTIISDPTCEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLGDDTSLAL 466  
QY 481 TSALISIPCKGLGSLRTVNGVLSIGCLAMFVTMTLLYKKHKYKPIGNCNTRNVYKG 540  
Db 467 TSTLISVDPDRPASPLRMANSALLISVGCIAIFVTVISLLYKKHKYKPIENSPGNVYVS 526  
QY 541 KGLSVFLSHAKAPFSRGDRKDPQLQDK 568  
Db 527 KGLSVFLNRAKAVFPFGNQEKDPLLKQ 554  
RESULT 6  
US-09-850-716A-225  
: Sequence 225, Application US/09850716A  
: Patent No. US20020115139A1  
: GENERAL INFORMATION:  
: APPLICANT: Kalos, Michael D.  
: APPLICANT: McNeill, Patricia D.  
: APPLICANT: Retter, Marc W.  
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
: FILE REFERENCE: 210121.455C15  
: CURRENT APPLICATION NUMBER: US/09/850,716A  
: CURRENT FILING DATE: 2001-05-07  
: NUMBER OF SEQ ID NOS: 440  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 225  
: LENGTH: 560  
: TYPE: PRT  
: ORGANISM: Homo sapien  
US-09-850-716A-225

Query Match 70.7%; Score 2163; DB 10; Length 560;  
Best Local Similarity 69.4%; Pred. No. 9.8e-171;  
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;  
QY 1 MESLGGVFLVLLAAGLPQAAKRFVDVLGHEQYDPDHMRNNQLRGWSSDENEDQLYLP 60  
Db 1 MECLYFLGFLLLAARLPDAAKRFHDVLGNERP SAYMRHNLNGWSSDENEDWNEKLYP 60  
QY 61 VWRGEGRWKDSWEGRGVQAALTSDPALVGSNITFVNLVFPRCQKEDANGNIYVERNC 120  
Db 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPCQKEDANGNIYVERNC 120  
QY 121 RSLELASDPYVYNTTGADDEDWEDNTSQOHLRFPDGPFPFRPHGKKNWVYVHTL 180  
Db 121 RNEAGLSADPYVYNTWTAWSDESDGNGTQSHHNVFPDGPFPFRPHGWRWNFIYVHTL 180  
QY 181 GOYFQKLGCSARVSINTVNLTVGPQVMEVIVFRHGRAYIPISKVKDVIYDQIPFV 240  
Db 181 GOYFQKLGCSVRVSNTANVTLPQMEVTVYRRGRAYVPIAQVKDVIYVTDQIPFV 240  
QY 241 TMYQKNDNRNSDDETLRDLPIFDVLHDPShFLNYSIAISKWNGDNTGLFVSNHTLN 300  
Db 241 TMYQKNDNRNSDDETLRDLPIFDVLHDPShFLNYSIAISKWNGDNTGLFVSTNHTVN 300  
QY 301 HTVVLNGTFFNLTVTAVPQCPSPSPSSSTSPSPASSPTLSTPSPSLMPTGYKS 360  
Db 301 HTVVLNGTFFNLTVTAAPGCPPPPPPPRP-----SKPTPSLGPAGDNP 346  
QY 361 MELSDISNENCRINRYGYFRATITIVDGILEVNIIOVADVPIPTLPQDNLMDFIYTCCK 420  
Db 347 LELSRIPEDCOINRYGHFQATITIVEGILEVNIIOQTDVLMVPVPPESSLIDFVVTCQ 406  
QY 421 ATPTACTIISDPTCQIAONRVCSFVAVDELCLLSVRRAPNGSGTYCVNFTLGGDASLAL 480  
Db 407 SIPTVECTIISDPTCEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLGDDTSLAL 466  
QY 481 TSALISIPCKGLGSLRTVNGVLSIGCLAMFVTMTLLYKKHKYKPIGNCNTRNVYKG 540  
Db 467 TSTLISVDPDRPASPLRMANSALLISVGCIAIFVTVISLLYKKHKYKPIENSPGNVYVS 526

QY 541 KGLSVFLSHAKAPFSRGDRKDPQLQDK 568  
Db 527 KGLSVFLNRAKAVFPFGNQEKDPLLKQ 554  
RESULT 7  
US-09-897-778-225  
: Sequence 225, Application US/09897778  
: Patent No. US20020147143A1  
: GENERAL INFORMATION:  
: APPLICANT: Wang, TongLong  
: APPLICANT: Warnerakis, Margarita  
: APPLICANT: Fanger, Gary R.  
: APPLICANT: Vedvick, Thomas S.  
: APPLICANT: Carter, Darick  
: APPLICANT: Watanabe, Yoshihiro  
: APPLICANT: Henderson, Robert A.  
: APPLICANT: Peckham, David W.  
: APPLICANT: Fanger, Neil  
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
: FILE REFERENCE: 210121.455C16  
: CURRENT APPLICATION NUMBER: US/09/897,778  
: CURRENT FILING DATE: 2001-06-28  
: NUMBER OF SEQ ID NOS: 467  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 225  
: LENGTH: 560  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
US-09-897-778-225

Query Match 70.7%; Score 2163; DB 10; Length 560;  
Best Local Similarity 69.4%; Pred. No. 9.8e-171;  
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;  
QY 1 MESLGGVFLVLLAAGLPQAAKRFVDVLGHEQYDPDHMRNNQLRGWSSDENEDQLYLP 60  
Db 1 MECLYFLGFLLLAARLPDAAKRFHDVLGNERP SAYMRHNLNGWSSDENEDWNEKLYP 60  
QY 61 VWRGEGRWKDSWEGRGVQAALTSDPALVGSNITFVNLVFPRCQKEDANGNIYVERNC 120  
Db 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPCQKEDANGNIYVERNC 120  
QY 121 RSLELASDPYVYNTTGADDEDWEDNTSQOHLRFPDGPFPFRPHGKKNWVYVHTL 180  
Db 121 RNEAGLSADPYVYNTWTAWSDESDGNGTQSHHNVFPDGPFPFRPHGWRWNFIYVHTL 180  
QY 181 GOYFQKLGCSARVSINTVNLTVGPQVMEVIVFRHGRAYIPISKVKDVIYDQIPFV 240  
Db 181 GOYFQKLGCSVRVSNTANVTLPQMEVTVYRRGRAYVPIAQVKDVIYVTDQIPFV 240  
QY 241 TMYQKNDNRNSDDETLRDLPIFDVLHDPShFLNYSIAISKWNGDNTGLFVSNHTLN 300  
Db 241 TMYQKNDNRNSDDETLRDLPIFDVLHDPShFLNYSIAISKWNGDNTGLFVSTNHTVN 300  
QY 301 HTVVLNGTFFNLTVTAVPQCPSPSPSSSTSPSPASSPTLSTPSPSLMPTGYKS 360  
Db 301 HTVVLNGTFFNLTVTAAPGCPPPPPPPRP-----SKPTPSLGPAGDNP 346  
QY 361 MELSDISNENCRINRYGYFRATITIVDGILEVNIIOVADVPIPTLPQDNLMDFIYTCCK 420  
Db 347 LELSRIPEDCOINRYGHFQATITIVEGILEVNIIOQTDVLMVPVPPESSLIDFVVTCQ 406  
QY 421 ATPTACTIISDPTCQIAONRVCSFVAVDELCLLSVRRAPNGSGTYCVNFTLGGDASLAL 480  
Db 407 SIPTVECTIISDPTCEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLGDDTSLAL 466  
QY 481 TSALISIPCKGLGSLRTVNGVLSIGCLAMFVTMTLLYKKHKYKPIGNCNTRNVYKG 540  
Db 467 TSTLISVDPDRPASPLRMANSALLISVGCIAIFVTVISLLYKKHKYKPIENSPGNVYVS 526

QY 541 KGLSVFLSHAKAPSRGDRKDPDLQDK 568  
 Db 527 KGLSVFLSHAKAPSRGDRKDPDLQDK 554

RESULT 8

US-09-943-075A-6  
 ; Sequence 6, Application US/09943075A  
 ; Patent No. US20020151486A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Popoff, Steven N.  
 ; APPLICANT: Safado, Faye F.  
 ; APPLICANT: Owen, Thomas A.  
 ; APPLICANT: Smock, Steven L.  
 ; TITLE OF INVENTION: Osteoactivin protein and Nucleic Acids Encoding the Same,  
 ; FILE REFERENCE: 71369,262  
 ; CURRENT APPLICATION NUMBER: US/09/943,075A  
 ; CURRENT FILING DATE: 2001-08-30  
 ; PRIOR APPLICATION NUMBER: US 60/229,006  
 ; PRIOR FILING DATE: 2000-08-30  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 560  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-943-075A-6

Query Match 70.7%; Score 2163; DB 10; Length 560;  
 Best Local Similarity 69.4%; Pred. No. 9.8e-171;  
 Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

QY 1 MESLGVFLVLLAAGLPQAKRFQVLCGHEQYPPHMRNNOLRGWSSDENWDEQLYP 60  
 Db 1 MECLYFGLFLLAARLPDAKRFHVLGNRPSAYMREHNLGNWSSDENWDEQLYP 60  
 QY 61 VNRGEGRWKDSWGGRRVQAALTSDSPALVGSNITFVNLVFPCKQKEDANGNIYERNK 120  
 Db 61 VNRGDMRWKSWGGRRVQAALTSDSPALVGSNITFVNLVFPCKQKEDANGNIYERNK 120  
 QY 121 RSDLELADPVVYNTTGCADDEWEDNTSOGHLRFQDCKPGRPHGRKKNFVVFHTL 180  
 Db 121 RNEAGLSADPVVYNTTGCADDEWEDNTSOGHLRFQDCKPGRPHGRKKNFVVFHTL 180  
 QY 181 GYFQKLGCSRVSVNTANVTLPQLMETVYRRHGRAYVPIAQVKDYYVVTQIPVFV 240  
 Db 181 GYFQKLGCSRVSVNTANVTLPQLMETVYRRHGRAYVPIAQVKDYYVVTQIPVFV 240  
 QY 241 TWQKNDNRSSDFTLRDLPIFFDVLTHDPSHFLNYSATSKWNGDNTGLFVSNHTLN 300  
 Db 241 TWQKNDNRSSDFTLRDLPIFFDVLTHDPSHFLNYSATSKWNGDNTGLFVSNHTLN 300  
 QY 301 HTYVLNGTFNLTAVQAPGCPSPPTSPSSSTSPASSPPTLSTPSPSLMPTGYKS 360  
 Db 301 HTYVLNGTFNLTAVQAPGCPSPPTSPSSSTSPASSPPTLSTPSPSLMPTGYKS 360  
 QY 361 MELSDISNENCRINRYGYFRATITVIGILEVNIQVADVPITLPQDNLMDFTVCKG 420  
 Db 361 MELSDISNENCRINRYGYFRATITVIGILEVNIQVADVPITLPQDNLMDFTVCKG 420  
 QY 421 ATPTACTIISDPTCQIAQRVCSVAVDELCLLSVRRAFNGSGTYGVNFTLGDASIAL 480  
 Db 421 ATPTACTIISDPTCQIAQRVCSVAVDELCLLSVRRAFNGSGTYGVNFTLGDASIAL 480  
 QY 481 TSALISIFKDLGSLPRTVNGVLSIGCLAMFTVMTILLYKKHKTYPKIGNCTRNKVG 540  
 Db 481 TSALISIFKDLGSLPRTVNGVLSIGCLAMFTVMTILLYKKHKTYPKIGNCTRNKVG 540  
 QY 541 KGLSVFLSHAKAPSRGDRKDPDLQDK 568  
 Db 527 KGLSVFLSHAKAPSRGDRKDPDLQDK 554

RESULT 9

US-10-227-881-42  
 ; Sequence 42, Application US/10227884  
 ; Publication No. US20030027988A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Gerritsen, Mary  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Philippe F.  
 ; APPLICANT: Watanabe, Colin L.  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3530P1C79  
 ; CURRENT APPLICATION NUMBER: US/10/227,884  
 ; CURRENT FILING DATE: 2002-08-26  
 ; PRIOR APPLICATION NUMBER: 10/119,480  
 ; PRIOR FILING DATE: 2002-04-09  
 ; PRIOR APPLICATION NUMBER: 60/059113  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/062287  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/063549  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/064103  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/069873  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/081819  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081955  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/082804  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/086392  
 ; PRIOR FILING DATE: 1998-05-22  
 ; PRIOR APPLICATION NUMBER: 60/089532  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089905  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090691  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/095302

```
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096146
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098544
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
```

```
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/125259
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 60/125775
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/126773
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/127887
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835
```

Query Match 70.6%; Score 2162; Db 9; Length 572;

Best Local Similarity 69.1%; Pred. No. 1.2e-170;

Matches 396; Conservative 78; Mismatches 87; Indels 12; Gaps 2;

Qy 1 MESLGVFLVFLLLAAGLPLQAAKFRDVLGHEQYDPDHRENNQLRGWSSDENEWDEQLYP 60

Db 1 MECLYVFLGLFLLAARLPDAAKFRHDVLGNRPSAYMRHNNQLNGWSSDNDWNEKLYP 60

Qy 61 VWRGGRWKNSWEGRGVOAALTSDSPALVGSNITFVNLVPPRCOKDANGNLYVERNC 120

Db 61 VWRGDMRWKNSWKGRVOAVLTSDSPALVGSNITFVNLVPPRCOKDANGNLYVERNC 120

Qy 121 RSDLEASDPYVYNWTTGADDEWEDNTSOGHLEFRPDGKPPRPHGKKKNFYVFFHTL 180

Db 121 RNEAGLSADPYVYNWTTAWSEDSGNGTQSHHNVFPDCKPPPHPGWRRWNFYVFFHTL 180



Oy 181 GYFQKLGOCARSINTVNLTVGPOVMEVIVFRBGRAYIPISKVKDYVYITDQIPFV 240  
Db 181 GYFQKLGRCSVRVNTVNTLGPOLMEVTVRRGRAYVPIAQKDVYVYVTVTQIPFV 240  
Oy 241 TMYOKNDRSSDETFELRDLPIFFDVLIHDPHSFLNYSATSKWNEGDNGLFVSNHNLN 300  
Db 241 TMYOKNDRSSDETFELRDLPIFFDVLIHDPHSFLNYSATSKWNEGDNGLFVSNHNLN 300  
Oy 301 HTYVLNCTFNFLTVQTVAPGCPSPTPSSSTSPSPASSPTLST-----PSPSLMP 355  
Db 301 HTYVLNCTFNFLTVQTVAPGCPSPTPSSSTSPSPASSPTLST-----PSPSLMP 355  
Oy 356 TGYKSMELSDISNENCRINRYGYFRATITVDGILEVNIIOVADVPITPLQPNLSLMDFI 415  
Db 356 TGYKSMELSDISNENCRINRYGYFRATITVDGILEVNIIOVADVPITPLQPNLSLMDFI 415  
Oy 416 VTCKGATPTACTIISDPTCOIAONRVCSVAVDELCLLSVRRFNGSGTCVNFLLGDD 475  
Db 416 VTCKGATPTACTIISDPTCOIAONRVCSVAVDELCLLSVRRFNGSGTCVNFLLGDD 475  
Oy 476 ASLATSALISIFGKDLGSPRLTVNGVLISIGCLAMFVMTVITLLYKHKTKYKIGNCTR 535  
Db 476 ASLATSALISIFGKDLGSPRLTVNGVLISIGCLAMFVMTVITLLYKHKTKYKIGNCTR 535  
Oy 536 NVYKGLSVFLSHAKAPSRGDKRDLQDK 568  
Db 536 NVYKGLSVFLSHAKAPSRGDKRDLQDK 568  
Oy 534 NVYKGLSVFLSHAKAPSRGDKRDLQDK 568  
Db 534 NVYKGLSVFLSHAKAPSRGDKRDLQDK 568

RESULT 10  
US-10-230-163-42  
; Sequence 42, Application US/10230163  
; Publication No. US20030036635A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3530PIC96  
; CURRENT APPLICATION NUMBER: US/10/230,163  
; PRIOR FILING DATE: 2002-08-28  
; PRIOR APPLICATION NUMBER: 10/119,480  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/084441  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/086392  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089905  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090691  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/095302  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/095318  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/095916  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/096146  
; PRIOR FILING DATE: 1998-08-11  
; PRIOR APPLICATION NUMBER: 60/096791  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 60/097986  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/098544  
; PRIOR FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: 60/099596  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099598  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099803  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099811  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099812  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099816  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100038  
; PRIOR FILING DATE: 1998-09-11  
; PRIOR APPLICATION NUMBER: 60/100385  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100390  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100627  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100848  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/100919  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/101477  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101738  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101741  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101786



```

; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C14
; CURRENT APPLICATION NUMBER: US/10/218,631
; PRIOR FILING DATE: 2002-08-12
; PRIOR FILING DATE: 2002-04-09
; PRIOR FILING DATE: 1997-10-17
; PRIOR FILING DATE: 1997-09-17
; PRIOR FILING DATE: 1997-10-28
; PRIOR FILING DATE: 1997-10-31
; PRIOR FILING DATE: 1997-12-17
; PRIOR FILING DATE: 1998-03-20
; PRIOR FILING DATE: 1998-03-25
; PRIOR FILING DATE: 1998-03-26
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 42
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-218-631-42

Query Match
Best Local Similarity 70.6%; Score 2162; DB 9; Length 572;
Matches 396; Conservative 78; Mismatches 87; Indels 12; Gaps 2;

QY 1 MESICGVLVFLLAAGLPQAARFRDVLGHEQYPDHMRNNOLRGWSSDENWDEQLYP 60
Db 1 MECLYFVLGFLLLAARLPDAARFRDVLGNERPSAYMREHNLNGWSSDENWNEKLYP 60

QY 61 VWRREGRWKDSWEGGRVQAALTSDSPALVGSNITFVNVLVPRCOKEDANGNTVYBNC 120
Db 61 VKRGDMRWKNSWGGVQAALTSDSPALVGSNITFVNVLVPRCOKEDANGNTVYBNC 120

QY 121 RSDLEASDPVYVNTTGADEDEDNTSGOHLRFDPGKPPRPHGRKKNFVYVHTL 180
Db 121 RNEAGLSADPVYVNTTGADEDEDNTSGOHLRFDPGKPPRPHGRKKNFVYVHTL 180

QY 181 GQYFQKLGCSARVSNITVLTGPOVMEVYVPRHGRAYIPISKVKOVYVITDQIPFV 240
Db 181 GQYFQKLGRCVSVSNITVLTGPOVMEVYVPRHGRAYIPISKVKOVYVITDQIPFV 240

QY 241 TMYQKNDNRNSDETFLRDLIPFDVLIHDPSPHFLANTYSAISKWNEFGDNTGLFVSNHNTLN 300
Db 241 TMEQKNDNRNSDETFLKDLIPFDVLIHDPSPHFLANTYSAISKWNEFGDNTGLFVSNHNTLN 300

QY 301 HTYVLTGTFNPLTVQTAVPDPCPTPSPSSSTSPSPASSPTLST-----PSPSLMP 355
Db 301 HTYVLTGTFNPLTVQTAVPDPCPTPSPSSSTSPSPASSPTLST-----PSPSLMP 355

QY 356 TGYKSMELSDISNENCRINRYGFRATITVDGILEVNTIQVADPIPTLPQDMSLMDFI 415
Db 356 TGDNPLELSRIPDENCQINRYGHQATITVEGILEVNTIQVADPIPTLPQDMSLMDFI 415

QY 416 VTCGATPTACTTIISDPTCOIAQNRVCSPVAVDELCLLSVRRFNGSGTCVNTPTLGD 475
Db 416 VTCGATPTACTTIISDPTCOIAQNRVCSPVAVDELCLLSVRRFNGSGTCVNTPTLGD 475

QY 476 ASLALTALISPGKDLGSLRTVNGVLISGCLAMFTVMTVILLYKKHKYKPGICNCTR 535
Db 476 ASLALTALISPGKDLGSLRTVNGVLISGCLAMFTVMTVILLYKKHKYKPGICNCTR 535

```

```

Db 474 TSLALTSTLISVPRDRPASPLRMANSALISVGCIAIFVTVISLVLVYKHKHKEYNPIENSPG 533
QY 536 NVYKCKGLSVFLSHAKAFPSRGDRKDPPLQDK 568
Db 534 NVYKCKGLSVFLSHAKAFPSRGDRKDPPLQDK 566

RESULT 12
US-10-230-338-42
; Sequence 42, Application US/10230338
; Publication No. US20030044934A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C92
; CURRENT APPLICATION NUMBER: US/10/230,338
; CURRENT FILING DATE: 2002-08-28
; PRIOR FILING DATE: 2002-04-09
; PRIOR FILING DATE: 1997-09-17
; PRIOR FILING DATE: 1997-10-17
; PRIOR FILING DATE: 1997-10-28
; PRIOR FILING DATE: 1997-10-31
; PRIOR FILING DATE: 1997-12-17
; PRIOR FILING DATE: 1998-03-20
; PRIOR FILING DATE: 1998-03-25
; PRIOR FILING DATE: 1998-03-26
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 42;
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-230-338-42

Query Match
Best Local Similarity 70.6%; Score 2162; DB 9; Length 572;
Matches 396; Conservative 78; Mismatches 87; Indels 12; Gaps 2;

QY 1 MESICGVLVFLLAAGLPQAARFRDVLGHEQYPDHMRNNOLRGWSSDENWDEQLYP 60
Db 1 MECLYFVLGFLLLAARLPDAARFRDVLGNERPSAYMREHNLNGWSSDENWNEKLYP 60

QY 61 VWRREGRWKDSWEGGRVQAALTSDSPALVGSNITFVNVLVPRCOKEDANGNTVYBNC 120
Db 61 VKRGDMRWKNSWGGVQAALTSDSPALVGSNITFVNVLVPRCOKEDANGNTVYBNC 120

QY 121 RSDLEASDPVYVNTTGADEDEDNTSGOHLRFDPGKPPRPHGRKKNFVYVHTL 180
Db 121 RNEAGLSADPVYVNTTGADEDEDNTSGOHLRFDPGKPPRPHGRKKNFVYVHTL 180

QY 181 GQYFQKLGCSARVSNITVLTGPOVMEVYVPRHGRAYIPISKVKOVYVITDQIPFV 240
Db 181 GQYFQKLGRCVSVSNITVLTGPOVMEVYVPRHGRAYIPISKVKOVYVITDQIPFV 240

```



```

Db 9  LLHLAVIGALLAVGATKVPKQDMLG-----VSRQLR-----TKAWNRQLYPEWT- 53
QY 65  GEGRWKDSWEGGRVQAALTSDSALVGSNITFVNLVFPFCQKEDANGNIVERN-CRSD 123
Db 54  -EQRLEDWRRGGVSLKVSNDGPTLIGANASFSIALNFPGSQKVLDPDGOVIVANNNTIING 112
QY 124  LELASDPVYVNWTTGADDEWEDNTSOGQHLRFPDGKPPRPHGRKKWKFVVFHTLQGY 183
Db 113  SQVMGGQPVYPOET--DDA-----CIPDGGPCPSGWSOKRSEFVVVWKTWGOY 159
QY 184  FOKLGQCSARVSINTVNLTVGPOMEVIVFRRHG-RAYIPISKVKDVYITDQIPFVTM 242
Db 160  WQVLGGPVSGLSICTGRAMLGTHMEVTVYHRRGSRVYPLAHSSSAFTITDQVPFVSIV 219
QY 243  YQKNDRNSDETEFLRDLPIFFDVLHDPHFNLNYSAISYKWNFGDNTGLFVSNHNTLNHT 302
Db 220  SOLRALDGGNKHFLRNQPLTFALQHDPSGLAEADLSYTWDFDSSGTLISRALVWVHT 279
QY 303  YVLNGTFNLTVOYAVP-----GPCP-----SNE----- 324
Db 280  YLEPGPVTAQVVLQAAIPLTSCGSPVPGTTDGHRTAEAPNTTAGOVPTTEVVGTTTPOQ 339
QY 325  SPTSPSSST-----SPSPASSPSTLSTPSPSLMP-----TG 357
Db 340  APTAEPSTTSVQVPTTEVISTAPVQMPTEAESTGMTPEKVPVSEVMGTTLAEMSTPEATG 399
QY 358  YKSMELSDI-----SNE----- 369
Db 400  MTPAEVSIIVLSGTTAAQVTTTETWETTARELPIPEPEGPDASSIMSTESITGSLGPLLD 459
QY 370  -----NCRINRYGYFRATITIVDGIENVIIQVADVPITPLQPDNSLMDFI 415
Db 460  GTATLRLVKRQVPLDCVLYRGVSFTLDIVQIESAEILQA-----VPSGEGD--AFELT 513
QY 416  VTCIKGATPEACTIISDPTCOIAQNRVCSVAVDELCLLSVRRAF-NGSGTYCVNFTLGD 474
Db 514  VSCGGGLPKKAEKWEISSPGCQPPAQRCLQPVLPSPACQVLVHLQILKGGSGTYCLNVSAD 573
QY 475  DASIALTSALISIPKDLG-SPLRTVNGVLISIGCLAMFVTMTVITYLKK 523
Db 574  TNSLAVVSTQIMPGQEAGLQVPLIVGIL-----LVLMVAVLASLIYRR 618

```

RESULT 15

US-09-862-260A-2  
 ; Sequence 2, Application US/09862260A  
 ; Patent No. US20020082217A1

GENERAL INFORMATION:

; APPLICANT: Nicolette, Charles A.  
 ; TITLE OF INVENTION: THERAPEUTIC ANTI-MELANOMA COMPOUNDS  
 ; FILE REFERENCE: 126881210200  
 ; CURRENT APPLICATION NUMBER: US/09/862,260A  
 ; PRIORITY FILING DATE: 2001-05-21  
 ; PRIOR APPLICATION NUMBER: 60/208,955  
 ; PRIOR FILING DATE: 2000-05-31  
 ; PRIOR APPLICATION NUMBER: 60/267,877  
 ; PRIOR FILING DATE: 2001-02-09  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 661

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-862-260A-2

Query Match 20.5%; Score 627.5; DB 10; Length 661;  
 Best Local Similarity 27.1%; Pred. No. 1.3e-43;  
 Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

```

QY 7  VLVPFLLAAGLPLQAAK--RFRDVLGHEQYDPHMRNNHOLRGWSSDENEDQLYPVVRR 64
Db 9  LLHLAVIGALLAVGATKVPKQDMLG-----VSRQLR-----TKAWNRQLYPEWT- 53

```

Search completed: April 7, 2003, 16:16:23  
 Job time : 22 secs

```

QY 65  GEGRWKDSWEGGRVQAALTSDSALVGSNITFVNLVFPFCQKEDANGNIVERN-CRSD 123
Db 54  -EQRLEDWRRGGVSLKVSNDGPTLIGANASFSIALNFPGSQKVLDPDGOVIVANNNTIING 112
QY 124  LELASDPVYVNWTTGADDEWEDNTSOGQHLRFPDGKPPRPHGRKKWKFVVFHTLQGY 183
Db 113  SQVMGGQPVYPOET--DDA-----CIPDGGPCPSGWSOKRSEFVVVWKTWGOY 159
QY 184  FOKLGQCSARVSINTVNLTVGPOMEVIVFRRHG-RAYIPISKVKDVYITDQIPFVTM 242
Db 160  WQVLGGPVSGLSICTGRAMLGTHMEVTVYHRRGSRVYPLAHSSSAFTITDQVPFVSIV 219
QY 243  YQKNDRNSDETEFLRDLPIFFDVLHDPHFNLNYSAISYKWNFGDNTGLFVSNHNTLNHT 302
Db 220  SOLRALDGGNKHFLRNQPLTFALQHDPSGLAEADLSYTWDFDSSGTLISRALVWVHT 279
QY 303  YVLNGTFNLTVOYAVP-----GPCP-----SNE----- 324
Db 280  YLEPGPVTAQVVLQAAIPLTSCGSPVPGTTDGHRTAEAPNTTAGOVPTTEVVGTTTPOQ 339
QY 325  SPTSPSSST-----SPSPASSPSTLSTPSPSLMP-----TG 357
Db 340  APTAEPSTTSVQVPTTEVISTAPVQMPTEAESTGMTPEKVPVSEVMGTTLAEMSTPEATG 399
QY 358  YKSMELSDI-----SNE----- 369
Db 400  MTPAEVSIIVLSGTTAAQVTTTETWETTARELPIPEPEGPDASSIMSTESITGSLGPLLD 459
QY 370  -----NCRINRYGYFRATITIVDGIENVIIQVADVPITPLQPDNSLMDFI 415
Db 460  GTATLRLVKRQVPLDCVLYRGVSFTLDIVQIESAEILQA-----VPSGEGD--AFELT 513
QY 416  VTCIKGATPEACTIISDPTCOIAQNRVCSVAVDELCLLSVRRAF-NGSGTYCVNFTLGD 474
Db 514  VSCGGGLPKKAEKWEISSPGCQPPAQRCLQPVLPSPACQVLVHLQILKGGSGTYCLNVSAD 573
QY 475  DASIALTSALISIPKDLG-SPLRTVNGVLISIGCLAMFVTMTVITYLKK 523
Db 574  TNSLAVVSTQIMPGQEAGLQVPLIVGIL-----LVLMVAVLASLIYRR 618

```



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 16:08:12 ; Search time 21 seconds  
(without alignments)  
2618.519 Million cell updates/sec

Title: US-09-943-075A-2

Perfect score: 3061

Sequence: 1 MESLGGVLLVLLAAGLPLO.....PFSRGDRKDPQLQDKPWML 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2163	70.7	560	2 138065	gene NMB protein -
2	629	20.5	662	2 138400	melanoma-associate
3	620	20.3	626	2 553871	Pmel 17 protein -
4	616	20.1	668	2 A41234	melanocyte-specifi
5	454	14.8	491	2 A49179	melanoma antigen h
6	133	4.3	888	2 A54280	cell differentiation
7	128.5	4.2	13055	2 T16580	hypothetical prote
8	126.5	4.1	611	2 T76211	hypothetical prote
9	125.5	4.1	351	2 S50754	hypothetical prote
10	123.5	4.0	279	2 T10361	hypothetical prote
11	123	4.0	473	2 S50755	hypothetical prote
12	121.5	4.0	422	2 T47901	hypothetical prote
13	121.5	4.0	795	2 T36288	endosperm specific
14	120.5	3.9	620	2 AG1598	probable penicilli
15	119.5	3.9	137	2 S37353	internalin like pr
16	119.5	3.9	273	2 C70551	nodulin (clone GmE
17	119	3.9	420	2 H84890	hypothetical prote
18	119	3.9	456	2 T17887	hypothetical prote
19	118.5	3.9	309	2 T17557	lysine/proline-ric
20	118	3.9	393	2 T33103	procyclin homolog
21	117.5	3.8	704	2 AE2107	lin-1 protein - Ca
22	116	3.8	1779	2 T31085	xylanase - Caldice
23	114	3.7	474	2 S15921	serine/threonine k
24	114	3.7	1331	2 A48954	protein TPX-VT3
25	113.5	3.7	446	2 T07907	mannan endo-1,4-be
26	111.5	3.6	221	2 T07176	hydroxyproline-ric
27	111.5	3.6	449	2 A24993	extensin homolog -
28	111.5	3.6	544	2 T17547	cellulase (EC 3.2.
29	111	3.6	4302	2 A38971	proline-rich prote
					polycystic kidney

```
30 110.5 3.6 681 2 T39076 serine rich pumili
31 110 3.6 168 2 T07640 pEARLI 1 protein -
32 110 3.6 1039 2 S02711 cellulase (EC 3.2.
33 108 3.5 360 2 S12850 protein TPX - Ther
34 108 3.5 560 2 AD2389 serine/threonine k
35 108 3.5 877 2 F90070 Clumping factor B
36 107.5 3.5 602 2 AD2067 hypothetical prote
37 107 3.5 507 2 S05542 hypothetical prote
38 107 3.5 542 2 I39540 chitinase (EC 3.2.
39 107 3.5 1036 2 T31673 N-acetylglucosamin
40 106.5 3.5 1000 2 C82630 serine proteinase
41 106 3.5 627 2 D75393 serine proteinase,
42 105.5 3.4 224 2 D72861 Acorif-91 protein -
43 105.5 3.4 287 2 S65765 chitinase (EC 3.2.
44 105 3.4 582 2 T07952 lectin-like protei
45 105 3.4 3020 2 A43932 mucin 2 precursor,
```

#### ALIGNMENTS

##### RESULT 1

138065  
gene NMB protein - human  
C:Species: Homo sapiens (man)  
C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 05-Nov-1999  
C:Accession: 138065  
R:Weternan, M.A.; Ajubi, N.; van Dinter, I.M.; Degen, W.G.; van Muijen, G.N.; Ruttler  
Int. J. Cancer 60, 73-81, 1995  
A:Title: nmb, a novel gene, is expressed in low-metastatic human melanoma cell lines  
A:Reference number: 138065; MUID:95113576; PMID:7814155  
A:Accession: 138065  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-560 <RES>  
A:Cross-references: EMBL:X76534; NID:9666042; PIDN:CAA54044.1; PID:9666043  
C:Genetics:  
A:Gene: GDB:NMB  
A:Cross-references: GDB:120237; OMIM:162340  
A:Map position: 15q22-15qter

Query Match 70.7%; Score 2163; DB 2; Length 560;

Best Local Similarity 69.4%; Pred. No. 1.7e-149;

Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

```
QY 1 MESLGGVLLVLLAAGLPLOAKRFQVLDHRENNQLRGWSSDENWDEQLYP 60
DB 1 MCCLYFFLGFLLLAARLPDAAKRFHDVLGNERPSAYMREHNQLNGWSSDENWNEKLYP 60
QY 61 VMRRGEGRKDSWEGRRVQAAATSDSPALVGSNITFFVNVLPFRCKEDANGNIVYERN 120
DB 61 VMKRGDMRKNSWGRVQAVLTSDSPALVGSNITFAVNLIFFPCOKEDANGNIVYERN 120
QY 121 RSDLEASDPVYVNTTGADEDEDNTSQGHLRFFDGKPPRPHGRKKNFVYVFTL 180
DB 121 RNEASADPTVYVNTTAMSESDGNGTSGSHHNVFPDGKPPHPCWRWNFVYVFTL 180
QY 181 GOYFKLGOCARSYINTVNLVGPQVMEVTVFRHGRVAPIPSKYKDVVITDQIPFV 240
DB 181 GOYFKLGRCRSVRVSVNTANVTLPQLMEVTVRRHGRVAPIAQVKDVVITDQIPFV 240
QY 241 TMYQNDNRSSDETFLRDLPIFFOVLIDHPSHFLNYSATSKWNFGDNTGLFVSNHHTLN 300
DB 241 TMFQNDNRSSDETFLKDLPIFMDVLIHPSHFLNYSTINYSKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVYNGTFNENLVQTAAGPCPCPTSPSSSTSPSPASSPTLTSTPSPSLMPTGYKS 360
DB 301 HTYVYNGTFNENLVQTAAGPCPCPTSPSSSTSPSPASSPTLTSTPSPSLMPTGYKS 360
QY 361 MELSDISNENCRINRYGFYRATITVDGILEVNIQVADVPIPTLPDNLMDFIYTCGK 420
DB 361 MELSDISNENCRINRYGFYRATITVDGILEVNIQVADVPIPTLPDNLMDFIYTCGK 420
QY 420 LELSTIPDENCQINRYGHFQATITVEGILEVNIQVADVPIPTLPDNLMDFIYTCGK 480
DB 420 LELSTIPDENCQINRYGHFQATITVEGILEVNIQVADVPIPTLPDNLMDFIYTCGK 480
```







QY 55 DEOLYVWRRGRGWKDSWEGGRVQAALTSQSPALVGSNITFVY-----NLVFPRCQKE-- 108  
|| : : : : : ||| : : : : : || : : : : :  
Db 2757 DENGIPLSQDGSLPTDN--QGNYLIVLTSETPTK-----TLPIDESGNVYVITPKPDGT 2809  
QY 109 ----DANGNIVYE-----RNCRSDEL-----ASDPYVYNWTTGADDEWEDNTS 149  
|| : : : : : ||| : : : : : || : : : : :  
Db 2810 PLATDSTGSEFVTDGTTIAKDDEGRPLGDKVEVLPDASGNVYVPTVSDQTLPTDGT 2869  
QY 150 QGOH-LRFDPDGKPPRPHGRKKWNVYVHTLGOYFQKLGOCSARVSINTVNLTVGQVM 208  
|| : : : : : ||| : : : : : || : : : : :  
Db 2870 KTYPIRTGPDGTPLP-----TDASGGVIGPD-G 2896  
QY 209 EVIVRRHGRAYTIPISKVDVIVITDQIPFVTMYQKKNRN-----SSDETEFLRLDPI-- 261  
|| : : : : : ||| : : : : : ||| : : : : :  
Db 2897 EIMPTDENG--IPLSK-----DGRP-----LPTDNDGNVYVSPDEETSKELPIDDS 2941  
QY 262 ----FFDVLHDSHFNLWYSAISKYKNFGDNGLFVSNHNTL----- 299  
|| : : : : : ||| : : : : : || : : : : :  
Db 2942 GNVYPIITKPDGTPLAT-----DSTGSEFVTDGTTIEKDDGKPLGPGQILST 2990  
QY 300 ----NHTYVNLGTFNFNLTVOT-----AVPGPCSPSPSPSSSTSPSPASSPSTLSTP 349  
|| : : : : : ||| : : : : : || : : : : :  
Db 2991 DASGNVYVPPGLDSQILTVDYVYKPIVTVIGPDGTALPTDASGAAGPDGTPISTDET 3050  
QY 350 SP-----SLMPTCYKSMELSDISNENCRNRYGYFRATITVDGILEVNIIOQADVPIPT 404  
|| : : : : : ||| : : : : : || : : : : :  
Db 3051 EPLKOGKSILPT-----DDYGNF---VFVVSQELPYD-----AEVOTPI 3086  
QY 405 LOPDNLMDPIVTCGATPEACTIISDPTCOIAQNRVCSPVAVDELCLLSVRRAFNGSG 464  
|| : : : : : ||| : : : : : || : : : : :  
Db 3087 TKPDGTL-----ATDSSG-NYVNDNGDIIEKDDGKPLGPDGEVLPT-----DGTG 3132  
QY 465 TYCVNLTGDDASLALTSALISIPGKDLGSLPTVNG 501  
|| : : : : : ||| : : : : : || : : : : :  
Db 3133 NFYIPATSDGE-----VIPDSDSKPLYTIRG 3160  
  
RESULT 8  
S76211  
hypotheical protein slr0442 - *Synechocystis* sp. (strain PCC 6803)  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76211  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Tas DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S76211  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: rNA  
A:Residues: 1-611 >X  
A:Cross-references: EMBL:D90914; GB:AB001339; NID:q1653477; PIDN:BAA18470.1; PID:d101  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
  
Query Match 4.1%; Score 126.5; DB 2; Length 611;  
Best Local Similarity 25.8%; Pred. No. 0.25;  
Matches 42; Conservative 23; Mismatches 51; Indels 47; Gaps 6;  
  
QY 283 WNEGDNTGLFVSNHNTLNIHTYVNLGTFNFNLTVOTAVGCPSPPTPSSSTSPSPASSP 342  
|| : : : : : ||| : : : : : || : : : : :  
Db 491 WSDIDSS--LADDDALTEPPITYTGSF---LTASSS-FSPSPSPSPSPSPSPSPSPSP 543  
QY 343 SPTLS-TPSPSLMPTCYKSMELSDISNENCRNRYGYFRATITVDGILEVNIIOQADV 401  
|| : : : : : ||| : : : : : || : : : : :  
Db 544 SP 567  
QY 402 IPTLPQDNLMDPIVTCGATPEACTIISDPTCOIAQNRVCS 444  
|| : : : : : ||| : : : : : || : : : : :  
Db 568 SPSPSPSPSPPTVTVNQVN---KKACDGLGGTYSGNNKTCT 607

[illegible]

RESULT 14  
AGI598  
Internalin like protein (LPXTG motif) [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
R:Accession: AGI598  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Blocker, H.;  
Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;  
Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Maier,  
O.K.C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Noland, J.A.; Voss, H.; Wehlund,  
A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AGI598  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-620 <GLA>  
A:Cross-references: GB:AL592022; PIDM:CAC96559.1; PID:g16413801; GSPDB:GN00178  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: linJ328

Search completed: April 7, 2003, 16:10:32  
Job time : 30 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 16:08:09 ; Search time 14 Seconds  
(without alignments)  
1694.604 Million cell updates/sec

Title: US-09-943-075A-2

Perfect score: 3061

Sequence: 1 MESLCLVFLVLLAAGLPLQ.....PFSRGDRKDPQLLDKPMWL 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2163	70.7	560	1 NMB_HUMAN	Q14956 homo sapien
2	1481.5	48.4	559	1 QNR_COTJA	Q90372 cuturnix co
3	640.5	20.9	762	1 P115_CHICK	Q98817 gallus gall
4	627.5	20.5	661	1 PM17_HUMAN	P40967 mus musculu
5	620	20.3	626	1 PM17_MOUSE	Q06154 bos taurus
6	454	14.8	491	1 PM17_BOVIN	P42523 dictyosteli
7	133	4.3	888	1 LAGC_DICDI	P54583 acidothermu
8	131.5	4.3	562	1 GUN1_ACICE	O10341 orgyia pseu
9	123.5	4.0	279	1 Y091_NPVOP	O06885 dictyosteli
10	120	3.9	544	1 GP10_DICDI	O05544 glycine max
11	119.5	3.9	137	1 N551_SOYBN	Q9F066 chlamydomon
12	118.5	3.9	555	1 GP1_CHLRE	P19275 thermoprote
13	114	3.7	474	1 VTP3_TTVIV	P22533 caldocellum
14	114	3.7	1331	1 MANB_CALSA	P98161 homo sapien
15	112.5	3.7	4303	1 PKD1_HUMAN	Q07984 cellulomona
16	111.5	3.6	449	1 GUNA_CELFI	Q92359 schizosacch
17	110.5	3.6	681	1 YDHE_SCHPO	O15294 homo sapien
18	110	3.6	1036	1 OGT1_HUMAN	P10474 c endogluca
19	110	3.6	1039	1 GUNB_HALSA	P19274 thermoprote
20	108	3.5	260	1 VTPX_TTVV1	P93329 medicago tr
21	107.5	3.5	268	1 NO20_MEDTR	P13692 enterococcu
22	107	3.5	516	1 P54_ENTFC	P56558 rattus norv
23	107	3.5	1036	1 OGT1_RAT	P41479 autographa
24	105.5	3.4	224	1 Y091_NPVAC	Q02817 homo sapien
25	105	3.4	5179	1 MUC2_HUMAN	P28191 vaccinia vi
26	104.5	3.4	281	1 VA04_VACCV	P22534 caldocellum
27	104.5	3.4	1742	1 GUNA_CALSA	P07359 homo sapien
28	103.5	3.4	626	1 GPBA_HUMAN	P18170 drosophila
29	103.5	3.4	950	1 DC12_DROME	P18169 drosophila
30	103.5	3.4	1123	1 DC11_DROME	P18171 drosophila
31	103.5	3.4	1589	1 DC13_DROME	Q9V493 homo sapien
32	103.5	3.4	2700	1 ZAN_HUMAN	P52549 human herpe
33	103	3.4	738	1 VU47_HSV6Z	

34	102	3.3	535	1 SPKC_SYNY3	P74745 synechocyst
35	102	3.3	551	1 POOX_RABIT	Q28645 oryctoiaagus
36	101.5	3.3	1009	1 YE68_METJA	Q58863 methanococc
37	101.5	3.3	1333	1 UTX_MOUSE	O70546 mus musculu
38	101.5	3.3	1569	1 YPJA_ECOLI	P52143 escherichia
39	101.5	3.3	1774	1 PSAS_PENPA	P22367 penicillium
40	100.5	3.3	792	1 PCAP_MOUSE	Q824h2 mus musculu
41	99.5	3.3	533	1 EGRI_MOUSE	P08046 mus musculu
42	99.5	3.3	1586	1 SN22_HUMAN	P31531 homo sapien
43	98.5	3.2	3317	1 CADN_RAT	P98365 rattus norv
44	98	3.2	682	1 AMPH_CHICK	P50478 gallus gall
45	98	3.2	734	1 G13B_DICDI	P34116 dictyosteli

#### ALIGNMENTS

#### RESULT 1

NMB_HUMAN	ID	NMB_HUMAN	STANDARD	PRT	560 AA.
AC	Q14956;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Putative transmembrane protein NMB precursor (Transmembrane				
DE	glycoprotein HGFIN).				
GN	GNMB OR NMB OR HGFIN.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Melanoma;				
RX	MEDLINE=95111576; PubMed=7814155;				
RA	Weterman M.A.J., Ajubi N., van Dinter I.M.R., Degen W.G.J.,				
RA	van Muljen G.N.P., Rutter D.J., Bloemers H.P.J.;				
RT	"NMB, a novel gene, is expressed in low-metastatic human melanoma				
RT	cell lines and xenografts."				
RL	Int. J. Cancer 60:73-81(1995).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Paripheral blood;				
RC	Qian J., Bandari P.S., Maloof P.B., Gascon P., Harrison J.S.,				
RA	Rameshwar P.;				
RT	Hematopoietic growth factor-inducible neurokinin-1 type (HGFIN) gene				
RT	encodes a novel transmembrane glycoprotein that is induced by				
RT	cytokines and neurotrophic factors in bone marrow (BM) stroma but is				
RT	constitutively expressed in differentiated cells."				
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.				
CC	-!- FUNCTION: COULD BE A MELANOGENIC ENZYME (BY SIMILARITY).				
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (potential).				
CC	-!- TISSUE SPECIFICITY: NOT RESTRICTED TO THE MELANOCTIC LINEAGE.				
CC	-!- DEVELOPMENTAL STAGE: EXPRESSION IN POORLY METASTATIC MELANOMA CELL				
CC	LINES; NO EXPRESSION IN HIGHLY METASTATIC MELANOMA CELL LINES.				
CC	-!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.				
CC	-!- SIMILARITY: CONTAINS 1 PKD DOMAIN.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
CC	EMBL: X76534; CAA54044.1; ..				
DR	EMBL: AF322909; NAG42839.1; ..				
DR	Genew; HGNC:4462; GPNMB.				
DR	MIM; 604368; ..				
DR	InterPro; IPR000601; PKD_domain.				
DR	Pfam; PF03801; PKD; 1.				
DR	SMART; SM00089; PKD; 1.				

PROSITE; PS50093; PKD; 1.  
Signal; Transmembrane; Glycoprotein; Polymorphism.  
KW SIGNAL 1 21  
FT CHAIN 22 560 PUTATIVE TRANSMEMBRANE PROTEIN NMB.  
FT DOMAIN 22 486 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 487 507 POTENTIAL.  
FT DOMAIN 508 560 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 240 327 PKD.  
FT DOMAIN 320 332 POLY-PRO.  
FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 249 249 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 455 455 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARIANT 195 S -> C (IN DBSNP:530436).  
FT FT /FTID=VAR\_012076.  
FT VARIANT 197 N -> H (IN DBSNP:530413).  
FT FT /FTID=VAR\_012077.  
FT SEQUENCE 560 AA; 62643 MW; 570035B48CCE3ECC CRC64;  
  
Query Match 70.7%; Score 2163; DB 1; Length 560;  
Best Local Similarity 69.4%; Pred. No. 2.6e-151;  
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;  
  
QY 1 MESLGVLPVLLAAGLPLOAKRFDVLGHEOYPDHRENNOLRGWSSDENDEOLYP 60  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 1 MECLYFLGLFLLAARPLDAARRHDLGNERSAYMRHNQLNGWSSDENDEWLKYL 60  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 61 VMRRGEGRKDSWEGSRVOAALTSIPALVGSNITFFVNLVPFCQEKEDANGNTVYERN 120  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 61 VKMRGDWRKNSWKGRVQAULTSDSPALVGSNITFAVNLIFFRCQEKEDANGNTVYEKN 120  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 121 RSLDELASDPYVYNWTGTADDEDWDNETSQOHLEFPDGKPPRPGRKKWNFYVPHTL 180  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 121 RNPAIGLSADPYVYNWTAMSEDSGNGTQSHNVNFPDGKPPRHPCWRNFIVVPHTL 180  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 181 GQYFKLGOCARSINTVNLVINGOVMEVIFVRHRGRAYTIPIISKVDVYVTQIPIFV 240  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 181 GQYFKQKLGRCSRVSVNTANVLTGPQLMEVTYVRHGGRAYPIAQVKDVYVVTQIPEVF 240  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 241 TMOKNDNRNSDETFLRDLPFEFDVLIHPDSHFNLNYSKWNFGDNTGLFVSNNHTLN 300  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 241 TFOKNDNRNSDETFLKDLPIMFDVLIHPDSHFNLNYSKWFSGDNTGLFVSTNHTVN 300  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 301 HTYVLNGTFNFLTAVTPGPCPTPSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTS 360  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 301 HTYVLNGTFSLNLTAKAAPGPCPPPMPRP-----SKPTPSLPGAGDNP 346  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 361 MELSDSNNCRNRYGYFRATITVDGLENIIOADVPIPTLPDNLMDLFIVCKG 420  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 347 LEUSRIDENCQNRYGHQAQITIVGILEVNIILQMTDVLMPWPBPSSSIDFVTCQG 406  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 421 ATPEACTIISDPTQTIAQRVCSPVAVDELCLLSVRRAFNCGSGTYCVNFTLGGDASIAL 480  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 407 SIPTVECTIISDPTCTEITONTVCSPVDDEMCLLVRRTRFNGSGTYCVNLFGDDTSLAL 466  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 481 TSALISPKDKLSPLRTVNGVLISIGCLAMPVMTVITLYKKIKTKYKPGNCRTNNVYG 540  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 467 TSWLSVDPDDPASPLRMANSALISVGCLAIFVTVISLIVYKKHKKEYNPENSGCNVRS 526  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
QY 541 KGLSVFLSHAKAPFSRGRDREKPDLLQK 568  
| | | | | | | | | | | | | | : | | | | | | | | : | | | | | | | |  
DB 527 KGLSVFLNRKAVFFRFGNQEKDPLLKNQ 554

## RESULT 2

```

QY 8 LVFLLLAAGLPLQAAKRFVDLHGEQVDPDHMRNNOLRGWSSDENEDPOLYVWRRGEG 67
DB 8 LALLPLPAEAVLCAAAARFQDVLNGRTAP-VTHKKIQQWSSDONKWKNEKLYPFWEEDNDP 66
QY 68 RKWDSWEGGRVQAAALTSDDSPALVGSNTFTFVNLVFPKQKEDANGNVYVRCRSDLELA 127
DB 67 RKWDCWKGKVTTLKLVTDSPALVGSNTFTFVNLVFPKQKEDANGNVYVRCRSDLELA 126
QY 128 SDPVYVNTTGADEDEDNEDSOGHLRFPDGPFRPHGRKKNFVYVFTLGOYFOKL 187
DB 127 ODQVYVNTTGADEDEDNEDSOGHLRFPDGPFRPHGRKKNFVYVFTLGOYFOKL 186
QY 188 GOCARSVINTVNLVGPQWVEIVFRRHGRARIPIKVKDQVYVTDQIPFVTVYQKND 247
DB 187 GRSSANFSVNTANITLKGHMAVSIYRRGHSTYVPIARASTYVTVVTKIPILVMSQKHD 246
QY 248 RNSDDETLRLDPIFFDVLHDPHPLNYSIAISKWKNFGDNTGLFVSNHNTLNHTYVNLG 307
DB 247 RNSDDETLRLDPIFFDVLHDPHPLNYSIAISKWKNFGDNTGLFVSNHNTLNHTYVNLG 306
QY 308 TFNFMNLVOTAVGPCSPSPSPSSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 367
DB 307 NFTNLVTAQAIPIPECKPVPPTP-----SLTPAVTTDASSNSDPSAPNEMAEDNP 357
QY 368 NENCINRYGYFRATITVDGILEVNIQVADVPITLQPDNSLMDFIPTCKGATPTAC 427
DB 368 NENCINRYGYFRATITVDGILEVNIQVADVPITLQPDNSLMDFIPTCKGATPTAC 427
QY 428 TIISDPTCOIAQNRVCSVAVDELCLLSVRRFNGSGTCVNTPLGDDASLALTSALISI 487
DB 428 TIISDPTCOIAQNRVCSVAVDELCLLSVRRFNGSGTCVNTPLGDDASLALTSALISI 487
QY 488 PGKDLGSLRVTVNGVLISIGCLAMFVTMTVILLYKKHKTYPICNCTRNVRVKGKLSVFL 547
DB 488 PGKDLGSLRVTVNGVLISIGCLAMFVTMTVILLYKKHKTYPICNCTRNVRVKGKLSVFL 547
QY 548 SHAKAPFRSGOREKPLDQKP 569
DB 548 SHAKAPFRSGOREKPLDQKP 569
QY 535 SNKFAFFPKSTERNPLLSKP 556
DB 535 SNKFAFFPKSTERNPLLSKP 556

RESULT 3
P115_CHICK
ID P115_CHICK STANDARD: PRT: 762 AA.
AC Q98917;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 115 kDa melanosomal matrix protein precursor.
GN MP115.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=white leghorn; TISSUE=Retinal pigment epithelium;
RX MEDLINE=92020667; PubMed=1924173;
RA Mochizuki M., Agata K., Eguchi G.;
RT "Complete sequence and expression of a cDNA encoding a chicken
RT 115-kDa melanosomal matrix protein.";
RL Pigment Cell Res. 4:41-47(1991).
RN [2]
RP CHARACTERIZATION.
RC STRAIN=white leghorn; TISSUE=Retinal pigment epithelium;
RX MEDLINE=88311098; PubMed=3409326;
RA Mochizuki M., Agata K., Kobayashi H., Yamamoto T.S., Eguchi G.;
RT "Expression of gene coding for a melanosomal matrix protein
RT transcriptionally regulated in the transdifferentiation of chick
RT embryo pigmented epithelial cells.";
RL Cell Differ. 24:67-74(1988).

```

```

CC -1- FUNCTION: MIGHT BE REQUIRED FOR POLYMERIZATION OF MELANIN ONTO THE
CC CORE STRUCTURE OF MELANOSOMES WITH ENZYMIC FUNCTION OF TYROSINASE.
CC -1- SUBCELLULAR LOCATION: ON THE FIBROUS MATRIX STRUCTURE OF THE
CC PREMELANOSOME.
CC -1- TISSUE SPECIFICITY: SPECIFIC TO PIGMENTED EPITHELIAL CELLS AND
CC MELANOCYTES. NOT EXPRESSED IN LENS, NEURAL RETINA, BRAIN, HEART,
CC GLIAD OR LIVER.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE REDIFFERENTIATION OF
CC PIGMENTED EPITHELIAL CELLS (PEC).
CC -1- PTM: GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO THE PHE1-17/NMB FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PKD DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D88348; BAA13589.1; -.
CC InterPro: IPR000601; PKD_domain.
CC Pfam: PF00801; PKD; 1.
CC SMART: SM00089; PKD; 1.
CC PROSITE: PS50093; PKD; 1.
CC Signal: Glycoprotein; Repeat.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 762 115 KDa MELANOSOMAL MATRIX PROTEIN.
CC DOMAIN 223 323 PKD.
CC DOMAIN 441 532 4 X 20-24 AA APPROXIMATE TANDEM REPEATS.
CC REPEAT 441 464 1.
CC REPEAT 465 488 2.
CC REPEAT 489 508 3.
CC REPEAT 509 532 4.
CC CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 762 AA; 77356 MW; 172C8DBD84FDC7C6 CRC64;
CC
Query Match 20.9%; Score 640.5; DB 1; Length 762;
Best Local Similarity 26.0%; Pred. No. 2.1e-39;
Matches 19; Conservative 90; Mismatches 194; Indels 261; Gaps 21;
QY 4 LCGVLVFLLLAAGLPLQAAKRFVDLHGEQVDPDHMRNNOLRG-----WS---SDENW 54
DB 3 LHGAIV--LLAALLALVTAQO-----RGGRSRGGVKGSAWGRPAFRSW 46
QY 55 DEQLYVWRRGGRKDSWEGGRVQAAALTSDDSPALVGSNTFTFVNLVFPKQKEDANGNI 114
DB 47 DTAFYRPMWEGTARQNDRCWGGDVTDFISNDAPTLVGARATFSTALRFPQTQTVLPDGRV 106
QY 115 VYFENC--RSDLELADSPVYVNTTGADEDEDNEDSOGHLRFPDGPFRPHGRKKNW 172
DB 107 VNSCNCVTNCTRNLOGDPV-----YPEOLAEKSGVDFPDQPPRSNWKGRGR 154
QY 173 FVYVFTLGOYFOKLQKCSARVSNITVNLVGPQWVEIVFRRHGR-AYIPISKVKDQVYV 231
DB 155 FVYVFTLGOYFOKLQKCSARVSNITVNLVGPQWVEIVFRRHGR-AYIPISKVKDQVYV 231
QY 232 ITDQIPFVTVYQKNDNRNSDDETLRLDPIFFDVLHDPHPLNYSIAISKWKNFGDNTGL 291
DB 215 ITDQIPFVTVYQKNDNRNSDDETLRLDPIFFDVLHDPHPLNYSIAISKWKNFGDNTGL 291
QY 292 FVSNHNTLNHTYVNLVGTAVP-----GPCPS----- 325
DB 275 LISRPTVTHVTLQAGSFAARLVLAQALPLSSCGTSAPPVVDPTTGPVPSLGTATPVG 334
QY 326 PTPS-----PS-----SSTSPSPASS-----PSTLS 347
DB 335 PTPS-----PS-----SSTSPSPASS-----PSTLS 347

```

```

QY 348 T-----PSP-----SLMPT----- 356
DB 395 TAVADAAAAGTDPATDPLPPTSVSSGGDAPGTAVPTAEGSVAGVGTADVAATPGATA 454
QY 357 -----GYKSMELSD----- 365
DB 455 ADVAVDTAGTGDGAVGPTAAATAESTADPTAGATGDGAVGATAGTGDGAV 514
QY 366 -----ISN-----ENCRINRYGYFRA 381
DB 515 GPTAAATAESTADPTAGATVSSGSATAGATAEPLLLVKKRAPEAEPCTGCVLYRYGTST 574
QY 382 TITIVDGIENVITIOADVPIPTLPQNSLMDFLVTCKGATPTEACTIISDPTCOIAQNR 441
DB 575 ELNIVQGIESTVAIVQV--VPAAPREGSNS-VEIUTVTCESGLPREVCTVVADAECTAQQ 631
QY 442 VCSPVAVDELCLLSVRRAFNGSGYCVNFTLGGDASLAITS---ALISIPGRD-----IG 493
DB 632 TCSAVAPAPGQQLVRODFNOSGLYCLNVSLANGNGLAVASTHVAVGSIPSRQWHHAHRG 691
QY 494 SPLRTVNGVLISIGCL 509
DB 692 AALGTAHRCSGHRCL 707

RESULT 4
PM17_HUMAN
AC P40567; Q16565; Q14817; Q12763; Q14448;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanocyte protein Pmel 17 precursor (Melanocyte lineage-specific
DE antigen gp100) (Melanoma-associated ME20 antigen) (ME20M/ME20S)
DE (ME20-M/ME20-S) (95 kDa melanocyte-specific secreted glycoprotein).
GN SILV OR Pmel17 OR D1253E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=92021023; PubMed=1924386;
RA Kwon B.S., Chintamaneni C., Kozak C.A., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Barton D., Francke U., Kobayashi Y.,
RA Kim K.-K.;
RT "A melanocyte-specific gene, Pmel 17, maps near the silver coat color
RT locus on mouse chromosome 10 and is in a syntenic region on human
RT chromosome 12."
RL Proc. Natl. Acad. Sci. U.S.A. 88:9228-9232(1991).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=94327568; PubMed=7519602;
RA Adema G.J., de Boer A.J., Vogel A.M., Loenen W.A., Fiedor C.G.;
RT "Molecular characterization of the melanocyte lineage-specific
RT antigen gp100."
RL J. Biol. Chem. 269:20126-20133(1994).
[3]
SEQUENCE FROM N.A.
RX MEDLINE=96154052; PubMed=8592076;
RA Ballin T., Lee S.T., Spritz R.A.;
RT "Genomic organization and sequence of D12S53E (Pmel 17), the human
RT homologue of the mouse silver (sl) locus."
RL J. Invest. Dermatol. 106:24-27(1996).
[4]
SEQUENCE FROM N.A.; AND SEQUENCE OF 25-53.
RX MEDLINE=94235165; PubMed=8179825;
RA Marsh G.A., Marken J.S., Neubauer M., Aruffo A., Hellstroem I.,
RA Hellstroem K.E., Marquardt H.;
RT "Cloning and expression of the gene for the melanoma-associated ME20
RT antigen."
RL DNA Cell Biol. 13:87-95(1994).

```

```

RN [5]
RP SEQUENCE FROM N.A.
RA Kwon B.S., Kim K., Heng H.H., Shi X.M., Tsui L., Lee Z.H.,
RA Youn B., Pickard R.T.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
[6]
RN [6]
RP SEQUENCE FROM N.A.
RA Vogel A.;
RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COULD BE A MELANOGENIC ENZYME. COULD REPRESENT AN
CC ONCOFETAL SELF-ANTIGEN THAT IS NORMALLY EXPRESSED AT LOW LEVELS IN
CC QUIESCENT ADULT MELANOCYTES BUT OVEREXPRESSED BY PROLIFERATING
CC NEONATAL MELANOCYTES AND DURING TUMOR GROWTH. RELEASE OF THE
CC SOLUBLE FORM, ME20-S, COULD PROTECT TUMOR CELLS FROM ANTIBODY
CC MEDIATED IMMUNITY.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL). THERE
CC IS ALSO A SECRETED SOLUBLE FORM, ME20-S, PROBABLY PRODUCT OF
CC PROTEOLYTIC CLEAVAGE.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOMAS. SOME
CC EXPRESSION WAS FOUND IN DYSPLASTIC NEVI. NOT FOUND IN NORMAL
CC TISSUES NOR IN CARCINOMAS.
CC -1- SIMILARITY: BELONGS TO THE Pmel-17/NMB FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PKD DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M77348; AAC60121.1; -
CC EMBL; S73003; AAC60634.1; -
CC EMBL; U31799; AAB00386.1; -
CC EMBL; U31808; AAB00386.1; JOINED.
CC EMBL; U31807; AAB00386.1; JOINED.
CC EMBL; U31797; AAB00386.1; JOINED.
CC EMBL; U31798; AAB00386.1; JOINED.
CC EMBL; U01874; AAB18479.1; -
CC EMBL; U20093; AAB19181.1; -
CC EMBL; U19491; AAB19181.1; JOINED.
CC EMBL; M32295; AAA5930.1; -
CC Genew; HGNC:10880; SILV.
CC MIM: 155550; -
CC InterPro: IPR000601; PKD_domain.
CC Pfam; PF00801; PKD; 1.
CC SMART; SM00089; PKD; 1.
CC PROSITE; PS50093; PKD; 1.
CC Transmembrane; Glycoprotein; Signal; Melanin biosynthesis; Repeat;
CC Antigen.
CC SIGNAL 1 24
CC CHAIN 25 661
CC DOMAIN 25 595
CC TRANSMEM 596 616
CC DOMAIN 617 661
CC DOMAIN 255 292
CC DOMAIN 315 444
CC REPEAT 315 327
CC REPEAT 328 340
CC REPEAT 341 353
CC REPEAT 354 366
CC REPEAT 367 379
CC REPEAT 380 392
CC REPEAT 393 405
CC REPEAT 406 418
CC REPEAT 419 431
CC REPEAT 432 444
CC CARBOHYD 81 81
CC CARBOHYD 106 106
CC CARBOHYD 111 111
CC CARBOHYD 321 321
CC CARBOHYD 568 568
CC
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).

```





QY 218 -RAYIPISKVKVYVITDQIPFVITMYQKNDNRSSDDELRLDPIFFOVLHDPHSFLNY 276  
Db 194 SOSVPLAHASSTFTTDQVPFVSYSQALDGETKHLRNLPLFALQLHDPSCYLAE 253  
QY 277 SATSYKWNFGDNTGLFVSNHNLHNYVINGFNFLVQTAVP-----GPCSP----- 326  
Db 254 ADLSYTWDFDGTGLISRALDVTHTYLESSTVTAQVLAQAIPLVSCGSSPVPGTTDGY 313  
QY 327 -----TPSPSSSTSPS-----PASSPSTLS----- 347  
Db 314 MPTAEPAGTTSROGTTKVVGTTPGOMPTTOPSGTTVQMPTEVTATTSQMLISAVTD 373  
QY 348 -----TPS-PS-----LMPTGYKSMELSDISNE----- 369  
Db 374 TTLAEVSTTEGTTPTRPSGTTVAQATTTEGPDASPLPLTQSGTSGISPLDLDTDITML 433  
QY 370 -----NCRINRYGFRATITIVDGIILEVNIIOVADVPITPQDPSLMDFIIVCKGAT 422  
Db 434 VKROVPLDCVLYRYGFSALDIVOGIESAEILQA--VFP-----SEGDAFELTVSCOGGL 487  
QY 423 PTEACTIISDPTQIAQNRVCSPVAVDELCLLSVRRAF-NGSGTYGVNFTGLGDDASLALT 481  
Db 488 PKEACMDISPGCOPPAQRCLQSPVSPDCQLVLHQLVKGSGCTYCLNVSLADANSALVA 547  
QY 482 SALISIPKDLG---SPLRTVNGVLISIGCLAMFVTMTLLYKHKYKPIGNCNTRNV 538  
Db 548 STQLVVPQDGLGQAPL-----LVGILLVAVVLASLILGIDLRSAQFPKCHM--- 598  
QY 539 KGKGLSVFLSHAKAPFSGRDREXPLL 565  
Db 599 -----VALTAAPASGLRAGLGENSPLL 621

RESULT 6  
PM17\_BOVIN STANDARD; PRT; 491 AA.  
AC Q06154;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Melanocyte protein Pmel 17 (Retinal pigment epithelial-specific protein) (Fragment).  
GN SILV OR PMEL17 OR RPEL.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
OC NCBI\_TaxID=9913;  
RN []  
RP SEQUENCE FROM N.A.  
RC TISSUE=Retina;  
RX MEDLINE=93122163; PubMed=1478275;  
RA Kim R.Y., Wistow G.J.  
RT "The CDNA RPE1 and monoclonal antibody HMB-50 define gene products preferentially expressed in retinal pigment epithelium.";  
RL Exp. Eye Res. 55:657-662(1992).  
CC -!- FUNCTION: COULD BE A MELANOGENIC ENZYME (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: RETINAL PIGMENT EPITHELIUM.  
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 PKD DOMAIN.

This SWISS-prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M81193; AAA30419.1; .  
DR InterPro; IPR000601; PKD\_domain.  
DR Pfam; PF00801; PKD; 1.

SMART; SM00089; PKD; 1.  
DR PROSITE; PS00093; PKD; 1.  
KW Transmembrane; Glycoprotein; Melanin biosynthesis; Repeat.  
FT NON\_TER 1  
FT DOMAIN <1 423  
FT TRANSMEM 424 444  
FT DOMAIN 445 491  
FT DOMAIN 60 150  
FT DOMAIN 148 256  
FT REPEAT 148 160  
FT REPEAT 161 173  
FT REPEAT 174 186  
FT REPEAT 187 199  
FT REPEAT 200 212  
FT REPEAT 213 225  
FT REPEAT 232 243  
FT REPEAT 244 256  
FT DOMAIN 304 394  
FT CARBOHYD 269 269  
FT CARBOHYD 396 396  
SQ SEQUENCE 491 AA; 51669 MW; 2BFE3DFBD397D6D CRC64;

Query Match 14.8%; Score 454; DB 1; Length 491;  
Best Local Similarity 26.1%; Pred. No. 5.8e-26;  
Matches 120; Conservative 71; Mismatches 137; Indels 132; Gaps 11;

QY 182 QYFQKLGQCSARVSNVNLTVGPQVMEVIVFRRHG-RAYIPISKVKVYVITDQIPFV 240  
Db 1 QYQVVLGPGVSGLSIGTDKAMILTYNMEVTVYHRRGQSYVPLAHSSSAFTITDQVPSV 60  
QY 241 TWQKNDNRSSDETFLRDLPIFFDLIHDPHSFLNYSAISYKWNFGDNTGLFVSNHNLN 300  
Db 61 SVSLOALDGRNKRFLRKQPLTFALQHDPSGLACADLSYTWDFGDSGTGLISRALTYT 120  
QY 301 HTYVLNGTFNFLTQVAVP-----GPCSP----- 326  
Db 121 HTYLESQVPTAQVVLQAAIPLTSCGSSPVPGTTDRHVTAEAPGTAGQVPTVEVMGTT 180  
QY 327 -----TPSPS 331  
Db 181 GOVPTAEAGTTVGVVPTTEDVGTPEQVATSKVLSTTPVEMPTAKATGRTPEVSTEPS 240  
QY 332 SS-----TPSPASSPSTLSTSP-----SLMPTGYKSMELSDISNE----- 369  
Db 241 GTTVTGTTPELVETTAGVSTPEPAGSNTSSPMPTAGTASLPLDDTATLVLEKQQA 300  
QY 370 --NCRINRYGFRATITIVDGIILEVNIIOVADVPITPQDPSLMDFIIVCKGATPEAC 427  
Db 301 PLDCVLYRYGFSLSLTDIVS-----IESAEILQAVSSEGDFAFELTVSCQGLPKAC 353  
QY 428 TIISDPTQIAQNRVCSPVAVDELCLLSVRRAP-NGSGTYGVNFTGLGDDASLALT 486  
Db 354 MDISSPGCQLPAQLCQPPPPSPACQLVLHQLVKGSGTYCLNVSLADANSALMWSTQLV 413  
QY 487 IPGKDLG---SPLRTVNGVLISIGCLAMFVTMTLLYK 523  
Db 414 MPQEAGLRQAPLEV--GILLVLTAL-----LLASLIYR 446

RESULT 7  
LAGC\_DICDI STANDARD; PRT; 888 AA.  
ID LAGC\_DICDI  
AC P42573;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Loose aggregate C protein precursor.  
GN LAGC.  
OS Dictyostellum discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.

```

Db 147 --IFGPNFLINSIDVKNPFVVKGNFSDPS--FDCNNITVTPP---PGSGKFRLYYDGTG 199
QY 320 PGFCPSPTPSPSSSTSPSPASSPTLSTPS---PSLAPTGYKSM-----LSDISNENC 371
      || || || || || || || || || || || || || || || || || || || || ||
Db 200 DNEVPFVESPIISSVSDSSKQIITINGDNFTDKDLVKVSPFDGIDQPNFIISVNHKQI 259
      || || || || || || || || || || || || || || || || || || || || ||
QY 372 RNRV-----GYFRATITVDGILEVNIQVADVPIPTLOPNSLMDFTVCKGATPEA 426
      :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 260 QVNNVRNDVPGPMSVNNITVDEVSIEKNNHCFPAITSISSVSNHLGDLVITIGBEKLSST 319
      || || || || || || || || || || || || || || || || || || || || ||
QY 427 CTIISDPTCOIAQNRVCSPVADEL-CLLSVRRAFNGSCTYCVNETLGDASLALTSALI 485
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 320 LNVYLTPTSTITGDKVYIIKSTTTELECKDLANELGGKNLPVNNVFGCDSTSPNGVSFTY 379
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 486 SIPGKDLGSLPTFNVGLISTG 507
      || || || || || || || || || || || || || || || || || || || || ||
Db 380 NIPTLSRGS---YSNGIVTLIG 398
      || || || || || || || || || || || || || || || || || || || || ||

```

GN1	AC1CE	STANDARD;	PRT;	562 AA.
AC	P54583;			
DT	01-OCT-1996	(Rel. 34, Created)		
DT	01-OCT-1996	(Rel. 34, Last sequence update)		
DT	01-OCT-1996	(Rel. 34, Last annotation update)		
DE	Endoglucanase E1 precursor (BC 3.2.1.4) (endo-1,4-beta-glucanase E1)			
DE	(cellulase E1) (Endocellulase E1).			
OS	Acidothermus cellulolyticus.			
OC	Bacteria; Actinobacteria; Actinobacteriia (class); Actinobacteridae;			
OC	Actinomycetales; Frankineae; Acidothermaceae; Acidothermus.			
NCBI	_TaxID=28049;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-ATCC 43068 / 11B;			
RC	Laymon R.A., Himmel M.E., Thomas S.R.;			
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 42-398.			
RX	MEDLINE=96346058; PubMed=8719855;			
KA	Sakon J., Adney W.S., Himmel M.E., Thomas S.R., Kaplus P.A.;			
RT	"Crystal structure of thermostable family 5 endocellulase E1 from			
RT	Acidothermus cellulolyticus in complex with cellobiose."			
RL	Biochemistry 35:10648-10660(1996).			
CC	-1- FUNCTION: THERMOSTABLE ENZYME WITH AN OPTIMAL TEMPERATURE OF 81			
CC	DEGREES CELSIUS. HAS A VERY HIGH SPECIFIC ACTIVITY ON			
CC	CARBOXYMETHYLCELULOSE.			
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic			
CC	linkages in cellulose.			
CC	-1- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL			
CC	HYDROLASES).			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			

the European Bioinformatics Institute. There are no restrictions on the use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.lsb-sib.ch/announcement/> or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

-----

CC EMBL; U33212; AAN75477.1; -  
CC PDB; 1ECE; 14-OCT-96.  
CC InterPro: IPR001919; Bac\_celose-bind.  
DR InterPro: IPR001347; GH\_5.  
DR Pfam: PFC0150; cellulase; 1.  
DR Pfam: PFC0553; CBM\_2; 1.  
DR PROSITE: PS00659; GLYCOSYL\_HYDROL\_F5; 1.  
DR Cellulose degradation; Hydrolase; Glycosidase; Signal; 3D-structure.  
KW SIGNAL  
FT SIGNAL 1 41  
FT CHAIN 42 562 ENDOGLUCANASE E1.  
FT DOMAIN 42 400 CATALYTIC.  
FT DOMAIN 401 461 PRO/SER/THR-RICH (LINKER).  
FT DOMAIN 452 562 CELLULOSE-BINDING (BY SIMILARITY).  
FT FT







```

FT  CONFLICT      338      338      T -> P (IN REF. 2).
FT  CONFLICT      340      346      TPTPTPT -> ROHQHRO (IN REF. 2).
SQ  SEQUENCE      1331 AA: 146892 MW: FFBCA51BB8D8F0E0 CRC64;

Query Match          3.7%; Score 114; DB 1; Length 1331;
Best Local Similarity 24.5%; Pred. No. 2;
Matches 27; Conservative 20; Mismatches 17; Indels 46; Gaps 4;

QY  317 TAVPGPCPTSPSSSTSPASSPSPTLS---TPSPSLMPTGYKSMELSDISNENCR 373
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  732 TVPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 773
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  374 NRYCYPRATITVDGILEVNIQVADVPITPLDPSNLMDFIV-TCKGAT 422
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  774 -----TVT-----PLPTISPSVVEITINTAGRT 799
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
PKD1_HUMAN STANDARD; PRT; 4303 AA.
AC P98161;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Polycystin precursor (Autosomal dominant polycystic kidney disease protein 1).
DE PKD1.
GN PKD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE-95254638; PubMed-7736581;
RA Gluecksmann-Kuls M.A., Taylor O., Woolf E.A., Bougueleret L.,
RA Deng N., Alperin G.D., Iris F., Hawkins F., Munro C., Lakey N.,
RA Dayk G., Schneider M.C., Geng L., Zhang F., Zhao Z., Torosian S.,
RA Reiders S.T., Bork P., Fohlschmidt M., Loehning C., Kraus B.,
RA Nowicka U., Leung A.L.S., Frischau A.-M.;
RA "Polycystic kidney disease: the complete structure of the PKD1 gene
    and its protein.";
    Cell 81:289-298(1995).
RN [2]
RP SEQUENCE OF 2769-4303 FROM N.A.
RP MEDLINE-94273192; PubMed-8004675;
RA Ward C.J., Peral B., Hughes J., Thomas S., Gamble V.,
RA Maccarthy A.B., Sloane-Stanley J., Buckle V.J., Kearney L.,
RA Higgs D.R., Ratcliffe P.J., Harris P.C., Roelise J.H.,
RA Spruit L.L., Saris J.J., Dauwerse H.G., Peters D.J.M.,
RA Breuning M.H., Nellist M., Brook-Carter P.T., Maheshwar M.M.,
RA Cordeiro I., Santos H., Cabral P., Sampson J.R., Janssen B.,
RA Hasselberg-Janssen A.L.W., van den Ouweland A.M.W., Eussen B.,
RA Verhoef S., Lindhout D., Halley D.J.J.;
RA "The polycystic kidney disease 1 gene encodes a 14 kb transcript and
    lies within a duplicated region on chromosome 16. The European
    Polycystic Kidney Disease Consortium.";
    Cell 77:881-894(1994).
RN [3]
RP STRUCTURE BY NMR OF 275-354.
RP MEDLINE-99107746; PubMed-9889186;
RA Bycroft M., Bateman A., Clarke J., Hamill S.J., Sandford R.,
RA Thomas R.L., Chothia C.;
RA "The structure of a PKD domain from polycystin-1: implications for
    polycystic kidney disease.";
    EMBO J. 18:297-305(1999).
RN [4]
RP VARIANT ADPKD 3748-ARG-VAL-3752 DEL, AND VARIANT ASP-3632.
RA Peral B., San Millan J.L., Ong A.C.M., Gamble V., Ward C.J.,
RA Strong C., Harris P.C.;
RA "Screening the 3' region of the polycystic kidney disease 1 (PKD1)
    gene reveals six novel mutations.";
    Am. J. Hum. Genet. 58:86-96(1996).

```

```

RN [5]
RP VARIANTS ADPKD P-2993; R-3016 AND V-3511, AND VARIANTS M-3510 AND
    F-4190.
RX MEDLINE-97342914; PubMed-9199561;
RA Peral B., Gamble V., Strong C., Ong A.C.M., Sloane-Stanley J.,
RA Zerres K., Winears C.G., Harris P.C.;
RA "Identification of mutations in the duplicated region of the
    polycystic kidney disease 1 gene (PKD1) by a novel approach.";
    Am. J. Hum. Genet. 60:1399-1410(1997).
RN [6]
RP VARIANT ALA-4058.
RX MEDLINE-97295081; PubMed-9150733;
RA Constantinides R., Xenophontos S.L., Neophytou P., Nomura S.,
RA Pierides A., Constantinou-Deltas C.D.;
RA "New amino acid polymorphism, Ala/Val4058, in exon 45 of the
    polycystic kidney disease 1 gene: evolution of alleles.";
    Hum. Genet. 99:644-647(1997).
RN [7]
RP VARIANTS T-2760; P-2761; V-2763; T-2764; Q-2791; T-2826; L-3008 AND
    L-3064.
RX MEDLINE-97449169; PubMed-9285784;
RA Watnick F.J., Plontek K.B., Cordal T.M., Weber H., Gandolph M.A.,
RA Qian F., Lens X.M., Neumann H.P.H., Germino G.G.;
RA "An unusual pattern of mutation in the duplicated portion of PKD1 is
    revealed by use of a novel strategy for mutation detection.";
    Hum. Mol. Genet. 6:1473-1481(1997).
RN [8]
RP VARIANT ADPKD THR-3678.
RX MEDLINE-97403939; PubMed-9259200;
RA Turco A.E., Rossetti S., Bresin E., Englisch S., Corra S.,
RA Pignatti P.F.;
RA "Three novel mutations of the PKD1 gene in Italian families with
    autosomal dominant polycystic kidney disease.";
    Hum. Mutat. 10:164-167(1997).
RN [9]
RP VARIANT ADPKD ASP-4032, AND VARIANT VAL-4045.
RX MEDLINE-98180892; PubMed-9521593;
RA Daniells C., Maheshwar M.M., Lazarou L., Davies F., Coles G.,
RA Ravine D.;
RA "Novel and recurrent mutations in the PKD1 (polycystic kidney
    disease) gene.";
    Hum. Genet. 102:216-220(1998).
RN [10]
RP VARIANT ADPKD MET-3375.
RX MEDLINE-99118881; PubMed-9921908;
RA Koptides M., Constantinides R., Kyriakides G., Hadjigavriel M.,
RA Patsalis P.C., Pierides A., Deltas C.C.;
RA "Loss of heterozygosity in polycystic kidney disease with a missense
    mutation in the repeated region of PKD1.";
    Hum. Genet. 103:709-717(1998).
RN [11]
RP VARIANTS ADPKD L-324 AND S-845, AND VARIANTS R-1399 AND L-1786.
RX MEDLINE-99294580; PubMed-10364515;
RA Thomas R.L., McConnell R., Whittacker J., Kirkpatrick P., Bradley J.,
RA Sandford R.;
RA "Identification of mutations in the repeated part of the autosomal
    dominant polycystic kidney disease type 1 gene, PKD1, by long-range
    PCR.";
    Am. J. Hum. Genet. 65:39-49(1999).
RN [12]
RP VARIANTS ADPKD P-2392 AND F-2423, AND VARIANTS R-1399; Q-2548 AND
    R-2638.
RX MEDLINE-20046890; PubMed-10577909;
RA Watnick T., Phakdeekitcharoen B., Johnson A., Gandolph M., Wang M.,
RA Briefel G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.;
RA "Mutation detection of PKD1 identifies a novel mutation common to
    three families with aneurysms and/or very-early-onset disease.";
    Am. J. Hum. Genet. 65:1561-1571(1999).
RN [13]
RP VARIANTS ADPKD 3994-L--F-3996 DUPL; G-4136 AND C-4154, AND VARIANTS.
RX MEDLINE-20441957; PubMed-10987650;
RA Perrichot R.A., Mercier B., Simon P.M., Whebe B., Clodes J., Ferec C.;
RA "DGE screening of PKD1 gene reveals novel mutations in a large cohort

```

of 146 unrelated patients.\*;  
 Hum. Genet. 105:231-239(1999).  
 [14]  
 RP VARIANTS ADPKD 3748-R--V-3752 DEL AND L-4132 DEL, AND VARIANT V-4045.  
 RX MEDLINE-20112427; PubMed=10647901;  
 RA Afzal A.R., Hand M., Ternes-Ferreira E., Saggar-Malik A., Taylor R.,  
 RA Jeffery S.;  
 RT "Novel mutations in the 3 region of the polycystic kidney disease 1  
 RT (PKD1) gene.\*;  
 RL Hum. Genet. 105:648-653(1999).  
 [15]  
 RN VARIANTS ADPKD PRO-4225 AND TRP-4276.  
 RP MEDLINE-99217041; PubMed=10200984;  
 RA Badenas C., Torra R., San Millan J.L., Lucero L., Mila M.,  
 RA Estivill X., Barnell A.;  
 RT "Mutational analysis within the 3' region of the PKD1 gene.\*;  
 RL Kidney Int. 55:1225-1233(1999).  
 [16]  
 RN VARIANTS ADPKD MET-2250 AND TRP-2329, AND VARIANTS CYS-2379; LEU-3066;  
 RP VAL-3139 AND LEU-3193.  
 RX MEDLINE-20311156; PubMed=10854095;  
 RA Perrichot R., Mercier B., Quere I., Carre A., Simon P., Whebe B.,  
 RA Cledes J., Ferec C.;  
 RT "Novel mutations in the duplicated region of PKD1 gene.\*;  
 RL Eur. J. Hum. Genet. 8:353-359(2000).  
 [17]  
 RN VARIANTS ADPKD PRO-2921 AND MET-3375, AND VARIANT LEU-3066.  
 RP MEDLINE-20382887; PubMed=10923040;  
 RA Koplides M., Mean R., Demetriadou K., Constantinides R., Pierides A.,  
 RA Harris P.C., Deltas C.C.;  
 RT "Screening of the PKD1 duplicated region reveals multiple single  
 RT nucleotide polymorphisms and a de novo mutation in Hellenic  
 RT polycystic kidney disease families.\*;  
 RL Hum. Mutat. 16:176-176(2000).  
 [18]  
 RN VARIANTS ADPKD GUN-3719 AND PRO-3852, AND VARIANT VAL-4045.  
 RP MEDLINE-20514565; PubMed=11058904;  
 RA Agutari G., Savelli S., Garbo M., Bozza A., Augello G., Penolazzi L.,  
 RA De Paoli Vitali E., La Torre C., Cappelli G., Piva R., del Senno L.;  
 RT "Novel splicing and missense mutations in autosomal dominant  
 RT polycystic kidney disease 1 (PKD1) gene: expression of mutated  
 RT genes.\*;  
 RL Hum. Mutat. 16:444-445(2000).  
 [19]  
 RN VARIANTS ADPKD SER-1166; GLU-1956; CYS-2408 AND GLY-2442--2443 INS,  
 RP AND VARIANTS HIS-1995 AND ASN-2604.  
 RX MEDLINE-20467506; PubMed=11012875;  
 RA Phakdeekitcharoen B., Watnick T.J., Ahn C., Whang D.-Y., Burkhart B.,  
 RA Germino G.G.;  
 RT "Thirteen novel mutations of the replicated region of PKD1 in an Asian  
 RT population.\*;  
 RL Kidney Int. 58:1400-1412(2000).  
 [20]  
 RN VARIANTS ADPKD TRP-3753 AND ASN-3815.  
 RP MEDLINE-20275386; PubMed=10729710;  
 RX Kim U.K., Jin D.K., Ahn C., Shin J.H., Lee K.B., Kim S.H., Chae J.J.,  
 RA Hwang D.Y., Lee J.G., Namkoong Y., Lee C.C.;  
 RT "Novel mutations of the PKD1 gene in Korean patients with autosomal  
 RT dominant polycystic kidney disease.\*;  
 RL Mutat. Res. 432:39-45(2000).  
 [21]  
 RN VARIANTS ADPKD Q-13; F-75; C-139; 1992-F-T-1993 DELINS L; 2220-R--P-  
 RP 2224 DEL; D-2336; D-2752; 2762-ILMR-2765 DUPL; M-2768; K-2771; P-2816;  
 RP S-2858; 3012-T--Y-3017 DEL AND 3748-L--R-3752 DEL, AND VARIANTS S-  
 RP 2674; M-2708; T-2734; L-2735; C-2765; M-2782; R-2814; G-2888; I-2905;  
 RP D-2966 AND L-3066.  
 RX MEDLINE-21063179; PubMed=11115377;  
 RA Rossetti S., Strmecki L., Gamble V., Burton S., Sneddon V., Peral B.,  
 RA Roy S., Bakkaloglu A., Komel R., Winearls C.G., Harris P.C.;  
 RT "Mutation analysis of the entire PKD1 gene: genetic and diagnostic  
 RT implications.\*;  
 RL Am. J. Hum. Genet. 68:46-63(2001).  
 [22]

Query Match 3.7%; Score 112.5; DB 1; Length 4303;  
 Best Local Similarity 18.3%; Pred. No. 12;  
 Matches 100; Conservative 64; Mismatches 178; Indels 203; Gaps 22;  
 QY 77 RVQAALTSPALVGSNITFVVNLVFPKCKEDANGNIVYERNCRSDLELASDPYVYNWT 136  
 Db 1018 RMOGLQVSTVPALVSPNATLALTA-----GVIV-----DSAVEVA-----FLMT 1056  
 QY 137 TGADDEDWEDNTSQGHLEFPDCKPFRPHGRKKWNVVYFHTLQYFOKLGOCARVSI 196  
 Db 1057 FG-----DGEQALHQFQ---PYNESEFPV----- 1078  
 QY 197 NTVNLTVGQVMEIV-----FRHGRAVPIKSKVDYVITDQIPFVFTMYQKDRN 249  
 Db 1079 -----DPSVAQVLEHNVHTYAAQGEYLLTVLASNAFENLTQQVPSV-----RA 1124  
 QY 250 S-----SDETFRLDLPFFDVLIDHPSHFLNYSATSKWNEGDMTGLFVSNHTLNH 301  
 Db 1125 SLPSSVAVGSDGVLVAGRPVTF---YHPPLSPGGVLYTDFDGDGSPVLTQSQAANH 1179  
 QY 302 TYVLNGTFFENLTVQTAVPGPCP-----SPTSPSSSTSPS 337  
 Db 1180 TYASRGTYHVRLEVNNTVSGAAAQADRVFELRGLSVDMSLAVEOGAPV---VVSAAVQ 1236  
 QY 338 PASSPSPTLSTPSPSLMPTGYKSMELSDISNENCR-----NRYGYFRATITIVDGIIEV 392  
 Db 1237 TGDNITWTFDMGDGTVLSGPEATVEHVYLRQACTVTVGAGSPAGHLSRLHVLVFLV 1296  
 QY 393 NIQVADVPIPLQDPSNLMDFI-----VTCKGATPEACTIISDP 433  
 Db 1297 LRVEPA-ACIPT-OPDARLTAYVTGNPAHYLFDWTFDGSNTTTRGC-PTVTHNFTTRSG 1353  
 QY 434 TCQIA-----ONR-----VCSPVAV-----DELCLLS----- 455  
 Db 1354 TFLALVLSRRNRAHYFTSICVEPEGVNTLQPERQVQLGDEAWLVACAWPPEPVRYT 1413  
 QY 456 -----VRRAFNGSGTCVYVNFGLDGDASLALTSALISIFPKDLGSLPT 498  
 Db 1414 WDFGTEEAAPTARGPEVTFIYRDPGSLVTVTASNNISAANDSALVEQEPVLVTSIK- 1472  
 QY 499 VNGVL 503  
 Db 1473 VNGSL 1477

Search completed: April 7, 2003, 16:08:40  
 Job time : 23 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 16:08:12 ; Search time 93 seconds

(without alignments)  
1267.301 Million cell updates/sec

Title: US-09-943-075A-2

Perfect score: 3061

Sequence: 1 MESLGVLFVLLAAGLPQ.....PFSRGDRKDPQLQDKPWML 572

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

1: sp.archaea:\*

2: sp.bacteria:\*

3: sp.fungi:\*

4: sp.human:\*

5: sp.invertebrate:\*

6: sp.mammal:\*

7: sp.mhc:\*

8: sp.organelle:\*

9: sp.phage:\*

10: sp.plant:\*

11: sp.rodent:\*

12: sp.virus:\*

13: sp.vertebrate:\*

14: sp.unclassified:\*

15: sp.rvirus:\*

16: sp.bacteriaph:\*

17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3061	100.0	572	11 Q9QZF6	Q9QZF6 rattus norv
2	2706	88.4	574	11 Q9QXA0	Q9QXA0 mus musculus
3	2700	88.2	574	11 Q99P91	Q99P91 mus musculus
4	711	23.2	206	4 Q96F58	Q96F58 homo sapien
5	634.5	20.7	626	11 Q9C2B2	Q9C2B2 mus musculus
6	398	13.0	461	6 Q97884	Q97884 equus caball
7	268	8.8	270	13 Q93391	Q93391 coturnix co
8	251	8.2	236	11 Q9QY67	Q9QY67 mus musculus
9	164	5.4	141	11 Q9QY70	Q9QY70 mus musculus
10	146	4.8	1698	17 Q8TPK7	Q8TPK7 methanosarc
11	139.5	4.6	1673	1 Q977V5	Q977V5 methanosarc
12	137.5	4.5	991	10 Q94C44	Q94C44 chlamydomon
13	134.5	4.4	955	10 Q94F92	Q94F92 chlamydomon
14	132	4.3	996	2 Q9A0H0	Q9A0H0 caldicellul
15	130.5	4.3	688	17 Q8TR88	Q8TR88 methanosarc
16	128.5	4.2	234	12 Q91GH4	Q91GH4 epiphyas po

17	128.5	4.2	13055	5	Q09165	Q09165 caenorhabdi
18	128	4.2	616	11	Q8R0X0	Q8R0X0 mus musculus
19	127	4.1	632	5	Q15742	Q15742 dictyosteli
20	127	4.1	752	5	Q9GUX3	Q9GUX3 dictyosteli
21	126.5	4.1	611	16	P74375	P74375 synechocyst
22	126	4.1	1817	17	Q8TI59	Q8TI59 methanosarc
23	126	4.1	3988	17	Q8TP21	Q8TP21 methanosarc
24	125.5	4.1	351	10	Q39492	Q39492 chlamydomon
25	123	4.0	473	10	Q39620	Q39620 chlamydomon
26	123	4.0	1426	2	Q9X3P6	Q9X3P6 caldicellul
27	123	4.0	2016	17	Q8TJS9	Q8TJS9 methanosarc
28	122	4.0	4283	11	Q9ERV0	Q9ERV0 rattus norv
29	121.5	4.0	422	10	Q9LZX4	Q9LZX4 arabidopsi
30	121.5	4.0	795	16	Q9X8T7	Q9X8T7 streptomyce
31	121.5	4.0	2566	17	Q8TSE7	Q8TSE7 methanosarc
32	121	4.0	2275	17	Q8TSE8	Q8TSE8 methanosarc
33	120.5	3.9	620	16	Q92C63	Q92C63 listeria in
34	120	3.9	323	10	Q9F845	Q9F845 arabidopsi
35	120	3.9	437	17	Q8ZXE3	Q8ZXE3 pyrobaculum
36	119.5	3.9	275	16	Q06170	Q06170 mycobacteri
37	119.5	3.9	906	17	Q8TPY9	Q8TPY9 methanosarc
38	119	3.9	420	10	Q22126	Q22126 arabidopsi
39	119	3.9	456	12	Q98436	Q98436 paramecium
40	118.5	3.9	309	12	Q89402	Q89402 paramecium
41	118	3.9	1046	11	Q81Y38	Q81Y38 mus musculus
42	113	3.9	1747	17	Q8TPY7	Q8TPY7 methanosarc
43	113	3.9	4578	13	Q42181	Q42181 fugu rubrip
44	117.5	3.8	704	16	Q8YUE0	Q8YUE0 anabaena sp
45	117.5	3.8	921	2	Q9L8L8	Q9L8L8 caldicellul

#### ALIGNMENTS

#### RESULT 1

Q9QZF6 PRELIMINARY; PRT; 572 AA.

AC Q9QZF6; 01-MAY-2000 (TREMREL. 13, Created)

DT 01-MAY-2000 (TREMREL. 13, Last sequence update)

DT 01-JUN-2002 (TREMREL. 21, Last annotation update)

DE Osteoactivin.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Xu J., Safadi F.F., Smock S.L., Rosenzweig A.B., Odgren P.R.,

RA Marks S.C. Jr., Popoff S.N., Owen T.A.;

RT "Cloning and characterization of a novel cDNA highly expressed in

RT osteopetrotic bone.";

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF184983; AAF03400.1; -

DR InterPro; IPR000601; PKD\_domain.

DR Pfam; PF00801; PKD; 1.

DR SMART; SM00089; PKD; 1.

DR PROSITE; PS00093; PKD; 1.

SQ SEQUENCE 572 AA; 63772 MW; CB939FB043EC57EB CRC64;

Query Match

Best Local Similarity 100.0%; Score 3061; DB 11; Length 572;

Mismatches 572; Conservative 0; Indels 0; Gaps 0;

QY 1 MESLGVLFVLLAAGLPQAKRFRDVLGHEQYPDHMRNNOLRGWSSDENWEDQLYP 60

DB 1 MESLGVLFVLLAAGLPQAKRFRDVLGHEQYPDHMRNNOLRGWSSDENWEDQLYP 60

QY 61 VWRREGRKWSWEGRGVQAALTSDSPALVGSNITTFVNLVFPKQKEDANGNIYVRC 120

DB 61 VWRREGRKWSWEGRGVQAALTSDSPALVGSNITTFVNLVFPKQKEDANGNIYVRC 120

QY 121 RSDLEASDPYVYNTTCADEDEDNTSQOHLRFPDGKFFPRPHGRKNWVYVFTL 180

Db	121	RSDELEASDPYYNWTGAEDEWENTSGOGLHAPPOGKPPPHGRKKWNVYVFHTL	180
Qy	181	GOYFOKLGOCARSINTVNLTVGPQVMEVIVFRHGRAYIPISKVDVVTIDQIPFV	240
Db	181	GOYFOKLGOCARSINTVNLTVGPQVMEVIVFRHGRAYIPISKVDVVTIDQIPFV	240
Qy	241	THYQKNDRNSDETFRLDLPFDVLHDPSHFLNYSALSYKWNFGDNTGLFVSNHNTLN	300
Db	241	THYQKNDRNSDETFRLDLPFDVLHDPSHFLNYSALSYKWNFGDNTGLFVSNHNTLN	300
Qy	301	HTYVLNGTFNFNLTVQTAVPGPCSPSTPSSSSTSPSPASSPSTLSTPSPSLMPTGYKS	360
Db	301	HTYVLNGTFNFNLTVQTAVPGPCSPSTPSSSSTSPSPASSPSTLSTPSPSLMPTGYKS	360
Qy	361	MELSDISNENCRINRYGYFRATITVDGILEVNIIOVADVPIPTLPQDPSLMDFI	420
Db	361	MELSDISNENCRINRYGYFRATITVDGILEVNIIOVADVPIPTLPQDPSLMDFI	420
Qy	421	ATPTEACTIISDPTCQIAQNRVCSPVAVDELCLISVRRAFNGSGTYCVNFTGLGDASL	480
Db	421	ATPTEACTIISDPTCQIAQNRVCSPVAVDELCLISVRRAFNGSGTYCVNFTGLGDASL	480
Qy	481	TSALISIPCKOLGSPLRVTNGVLISIGCLAMFVTWMTILLYKKHKTYPKIGNCTRNVVKG	540
Db	481	TSALISIPCKOLGSPLRVTNGVLISIGCLAMFVTWMTILLYKKHKTYPKIGNCTRNVVKG	540
Qy	541	KGLSVFLSHAKAPFSRGDRKDPLOQDKPWL	572
Db	541	KGLSVFLSHAKAPFSRGDRKDPLOQDKPWL	572

## RESULT 2

```

Q9QXA0
ID AC Q9QXA0 PRELIMINARY; PRT; 574 AA.
AC Q9QXA0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative transmembrane glycoprotein.
GN GPNMB OR NMB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bachner D., Schroder D., Gross G.;
RT "Isolation and developmental expression analysis of murine amb.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ251685; CAB65272.1; -.
DR MGD; MGI:1934765; Gpmdb.
DR InterPro: IPR000583; GATase_2.
DR InterPro: IPR000801; PKD_domain.
DR Pfam: PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN_1.
DR PROSITE; PS50093; PKD; 1.
KW Transmembrane.
SQ SEQUENCE 574 AA; 63681 MW; E0759D7626F0829A CRC64;
Query Match 88.4%; Score 2706; DB 11; Length 574;
Best Local similarity 89.0%; Pred. No. 6.2e-221;
Matches 508; Conservative 20; Mismatches 41; Indels 2; Gaps 1;

```

Qy	121	RSDELEASDPYYWNNTTCADDEDWNTSOCHLRPPDOGKPPRPHRGKKWNFYVHTL	180
Db	121	RNDLGITSLDHYNNWTAGADGDWEDTSRQHLRPPDRRPPRHGKKWSFYVHTL	180
Qy	181	GOYFQKLGQCSARVSINTVNLTVGCPQWMEYVIFRRHGRAYIPISKVKDYYVITDQIPFV	240
Db	181	GOYFQKLGRCARSINTVNLTVGCPQWMEYVYRRYGRAYIPISKVKDYYVITDQIPFV	240
Qy	241	TMVQKNDRNSDTEFLRDLPIFFDVLIIHDPSHFLANYSAISYKWNFGDNTGLFVSNHHTLN	300
Db	241	TMSQKKNDRLSDEIFLRLDPIFVDVLIIHDPSHFLNDSAISYKWNFGDNTGLFVSNHHTLN	300
Qy	301	HTVYLVNGTFENLTVQTAVPGCPSPSPS--PSSSTSPSPASSPSPSLTSPSPSLMPTGY	358
Db	301	HTVYLVNGTFENLTVQTAVPGCPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSLMPTGY	360
Qy	359	KSMELSDISNENCRNRYGYFRATITVDGILBWNIIQVADVIPITLQPNLSLMDFTVTC	418
Db	361	KSMELSDISNENCRNRYGYFRATITVEGILBWNIIQVADVIPITLQPNLSLMDFTVTC	420
Qy	419	KGATPTEACTIISDPTCOIAQNRVCSPVAVDELLCSLVSRRAPNMSGYVCYNFTLGGDASL	478
Db	421	KGATPMEACTIISDPTCOIAQNRVCSPVAVDGLLCSLVSRRAPNMSGYVCYNFTLGGDASL	480
Qy	479	ALTSALISIPGKDGSLPRTVNGVLSISIGCLAMFNTWITILLYKKHKTYPKIGNCNRNVY	538
Db	481	ALTSTUISIPGKDDPSLRPAVNGVLSISIGCLAVLNTWITILLYKKHKAYPKIGNCNRNVY	540
Qy	539	KGKGLSVFLSHAKAPFRRGDRKDPDLLQDKP	569
Db	541	KGKGLSVLLSHAKAPFRRGDRKDPDLLQDKP	571

### RESULT 3

ID	Q99p91	PRELIMINARY	PRT	574 AA
AC	Q99p91;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Dendritic cell-associated transmembrane protein (Glycoprotein			
DE	("transmembrane" nmb).			
GN	GNMB OR DCHIL.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RC	SPRAIN-BALB/C;			
RA	Shikano S., Bonkobara M., Zukas P.K., Ariizumi K.;			
RT	"Molecular Cloning of a Dendritic Cell-Associated Transmembrane			
RT	Protein, DC-HIL, that Promotes RGD-dependent Adhesion of Endothelial			
RT	Cells Through Recognition of Heparan Sulfate Proteoglycans.,"			
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.			
[2]				
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Salivary GLAND;			
RA	Strausberg R.;			
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF220254; AAA4240.1; -.			
DR	EMBL; BC026375; AAH26375.1; -.			
DR	MGI; MGI:1934765; Gpmb.			
DR	InterPro; IPR000583; GATase_2.			
DR	InterPro; IPR000601; PKD_domain.			
DR	InterPro; IPR002965; P_rich_extensn.			
DR	Pfam; PF00801; PKD; 1.			
DR	PRINTS; PR01217; PRICHEXTENS.			
DR	SMART; SM00089; FND; 1.			
DR	PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN_1.			
DR	PROSITE; PS50093; PKD; 1.			
DR	Transmembrane.			
SW	SEQUENCE 574 AA: 63674 MW: B6AE9AC27AE6ACD0 CRG64;			

Query Match	88.2%;	Score 2700;	DB 11;	Length 574;
Best Local Similarity	88.8%;	Pred. No. 2e-220;		
Matches 507;	Conservative 20;	Mismatches 42;	Indels 2;	Gaps 1;
QY 1	MESLCGVLFVLLAAGLPLQAARFRDVLGHEQYPOHMRNNQLRGWSSDENWDEQLYP 60			
Db	1	MESLCGVLFVLLAAGLPLQAARFRDVLGHEQYPOHMRNNQLRGWSSDENWDEHLYP 60		
QY 61	VWRGEGRWKDSWEGGRVQAAALSDSPALVGSNITFVNLVFPKCKEDANGNIYERNK 120			
Db	61	VWRGEGRWKDSWEGGRVQAAALSDSPALVGSNITFVNLVFPKCKEDANGNIYERNK 120		
QY 121	RSDELASDPYVYVNTTGADEDEMDNTSOGHRLRPDPKPPRPHGRKKWNVVVFHTL 180			
Db	121	RNEAGLSADPYVYVNTTGADEDEMDNTSOGHRLRPDPKPPRPHGRKKWNVVVFHTL 180		
QY 181	GOYFQKLGCSARVSNITVNLTVGPQVMEVIFVRHGRAYIPISKVKDYVITDQIPIFV 240			
Db	181	GOYFQKLGCSARVSNITVNLTVGPQVMEVIFVRHGRAYIPISKVKDYVITDQIPIFV 240		
QY 241	TYOKNDRNSDETFRLDLPFIFFDVLIHDPHFNLANSYKWNFGDNTGLFVSNHNTLN 300			
Db	241	TWOKNDRNSDETFRLDLPFIFFDVLIHDPHFNLANSYKWNFGDNTGLFVSNHNTLN 300		
QY 301	HTYVLNGTFNLTAVTAVGCPSPSPS--PSSSTSPSPASSPTLSTPSPSLMPTGY 358			
Db	301	HTYVLNGTFNLTAVTAVGCPSPSPS--PSSSTSPSPASSPTLSTPSPSLMPTGY 358		
QY 359	KSMELSDISNCRINRYGFRATITIVDGIENVLIQVADYPIPTLPQDNLMDFIYTC 418			
Db	361	KSMELSDISNCRINRYGFRATITIVDGIENVLIQVADYPIPTLPQDNLMDFTVTC 420		
QY 419	KGATPTEACTIISDPTQIAQNSVCSPVAVDELCLLSVRAENGSGTCVNFILGDDASL 478			
Db	421	KGATPTEACTIISDPTQIAQNSVCSPVAVDELCLLSVRAENGSGTCVNFILGDDASL 480		
QY 479	ALTSALISIPGKIDGLSPRTVNGVLISIGCLAMFVMTVITLLYKHKYKTKPNCNTRVY 538			
Db	481	ALTSTLISIPGKIDGLSPRTVNGVLISIGCLAMFVMTVITLLYKHKYKTKPNCNTRVY 540		
QY 539	KGKGLSVFLSHAKAFPSRGDREKPLLODKP 569			
Db	541	KGKGLSVFLSHAKAFPSRGDREKPLLODKP 571		
RESULT 4				
ID	Q96F58	PRELIMINARY;	PRT;	206 AA.
AC	Q96F58;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DE	Similar to glycoprotein (transmembrane) nmb.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=SKIN;			
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC011595; AAH11595.1; --			
KW	Transmembrane.			
SQ	SEQUENCE 206 AA; 23945 MW; 30CBDE6928D73FBD CRC64;			
Query Match	23.2%;	Score 711;	DB 4;	Length 206;
Best Local Similarity	69.2%;	Pred. No. 2.5e-52;		
Matches 128;	Conservative 23;	Mismatches 34;	Indels 0;	Gaps 0;
QY 1	MESLCGVLFVLLAAGLPLQAARFRDVLGHEQYPOHMRNNQLRGWSSDENWDEQLYP 60			

Db	1	MESLCGVLFVLLAAGLPLQAARFRDVLGHEQYPOHMRNNQLRGWSSDENWDEHLYP 60			
QY 61	VWRGEGRWKDSWEGGRVQAAALSDSPALVGSNITFVNLVFPKCKEDANGNIYERNK 120				
Db	61	VWRGEGRWKDSWEGGRVQAAALSDSPALVGSNITFVNLVFPKCKEDANGNIYERNK 120			
QY 121	RSDELASDPYVYVNTTGADEDEMDNTSOGHRLRPDPKPPRPHGRKKWNVVVFHTL 180				
Db	121	RNEAGLSADPYVYVNTTGADEDEMDNTSOGHRLRPDPKPPRPHGRKKWNVVVFHTL 180			
QY 181	GOYFQKLGCSARVSNITVNLTVGPQVMEVIFVRHGRAYIPISKVKDYVITDQIPIFV 240				
Db	181	GOYFQKLGCSARVSNITVNLTVGPQVMEVIFVRHGRAYIPISKVKDYVITDQIPIFV 240			
RESULT 5					
ID	Q9CZB2	PRELIMINARY;	PRT;	626 AA.	
AC	Q9CZB2;				
DT	01-JUN-2001 (TrEMBLrel. 17, Created)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)				
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)				
DE	N/A.				
GN	SI OR SI.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schriml J.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,				
RA	Hayashizaki Y.				
RT	"Functional annotation of a full-length mouse cDNA collection."				
RL	Nature 409:685-690(2001).				
DR	EMBL; AK012808; BAB28486.1; --				
DR	MGI; 98301; SI.				
DR	InterPro; IPR000601; PKD_domain.				
DR	Pfam; PF00801; PKD; 1.				
DR	SMART; SM00089; PKD; 1.				
DR	PROSITE; PS50093; PKD; 1.				
SQ	SEQUENCE 626 AA; 66301 MW; 7EC0A06C63212674 CRC64;				
Query Match	20.7%;	Score 634.5;	DB 11;	Length 626;	
Best Local Similarity	27.9%;	Pred. No. 3.7e-45;			
Matches 177;	Conservative 88;	Mismatches 191;	Indels 179;	Gaps 20;	
QY 54	WDEGLYPVWRGEGRWKDSWEGGRVQAAALSDSPALVGSNITFVNLVFPKCKEDANGN 113				
Db	43	WNRCLIPETVEQV--SNCWRGGQVSLRVWDGPTLVGNASFSIAUHFPGSRVLPDQ 100			
QY 114	IYVERNCRSDLEASDPYVYVNTTGADEDEMDNT----SQ----GQHLR-----FP 157				
Db	101	VI-----WANNIINGSQVGGQVYVQEPDQACVFP 132			
QY 158	DGKFFPRPHGRKKWNVVVFHTLGOYFQKLGCSARVSNITVNLTVGPQVMEVIFVRH 217				





```
Db 380 TSDVFL-SVPVQFT-----DLSENATKWNDFDGDG-----FTSIKQNPVHTYSEVGY 427
QY 310 NFNLTVQTAVPGPCSPTSPSSSTSPSPASSPSPT-----LSTPSPSLMPTGYKSMEL 363
Db 428 TVRLTVSNS-----NGTDSKLTAVNVVPRKGLSAPSYAYIANLSNTSVINTGNSLTT 481
QY 364 S-----DISNENCRI---NRYGYFRATITIVDGIIE--VNLIQVADVPIP---TL 405
Db 482 TVPVGIGPLGVAASPDGTRIVYVNSFNRYGTVSVIDTALNEVIAVDIGDKYSPGCIAY 541
QY 406 QPD-----NSLMDFTVTCGATPTEACTI-----TSDPCTQIAONRVCSPPVA 447
Db 542 TPDGKKLYVANRDIDGVSVIDTISINTVIATIPVGINPLGVAANPDGTVKVVVTVNRYSNVVS 601
QY 448 VDEL-----CLLSVRRAFNGSGTY---CVNFTLG--DDASLALTSALISI 487
Db 602 VIDFATNKVVATVKTGSGPCGCIIVNQ--EGTNLYVANCENNTISIIDTGS--NTATASV 656
QY 488 PGKDLGS-PLRTVNGVLISIGCLAMFV-----TWVTILYKKHKH-----525
Db 657 PA---GTWPM---GVAYSPDGKIVYVANERSNNVSVIDLATKTDIAAVKVGRCPYGIAV 709
QY 526 ----TYKPIGNCTRNVVYKGLSVF-----LSHAKAPFS 555
Db 710 TPDGTRVVAVANGCNQNLGKTVTSIIDTATNKVIATVKTGFS 750

RESULT 11
Q977V5
ID Q977V5 PRELIMINARY; PRT; 1673 AA.
AC Q977V5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Surface antigen.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-6;
RA Takagi J., Springer T.A.;
RT "Corrected sequence of surface antigen from Methanosarcina mazei S-6."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394229; AAK84029.1; -.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR000601; PKD_domain.
DR Pfam; PF01436; NHL; 11.
DR Pfam; PF00801; PKD; 12.
DR PROSITE; PS00093; PKD; 10.
SQ SEQUENCE 1673 AA; 175036 MW; 36B8CCB03AA7F11E CRC64;

Query Match 4.68; Score 139.5; DB 1; Length 1673;
Best Local Similarity 20.9%; Pred. No. 0.014;
Matches 103; Conservative 73; Mismatches 199; Indels 117; Gaps 25;

QY 165 PHGRKKWNVFVFTLGYFQKLGCSARVSINTVNLTVGPQVMEVIVFRHRGRAYIPIS 224
Db 251 PEGTK-----AYVTN-VDKYFTVSMIDGTNKITARPVGDPDAGIAVTPDGKKVYVALS 305
QY 225 KYKDYVI-TDQIPFIVTVQKNDRNSSDETFLRLDPI-----PFD-----VLIHD 269
Db 306 FCNTVSVIDTATNTITATMAVGKNPYASQG-FIGSIPVQVPYPSADPKSNITSGYIFLSE 364
QY 270 PSHF--LNYSAISKWNGDNTGLFVSNHNTLNHTYVLNGTFENLTVQTAVPGPCSPT 327
Db 365 PQVQFDLSKDATKWKDFDGDSS---SKQNPHTHTYSTGTIVRLTVSNS-----NGT 415
QY 328 PSPSSSTSPSPASSPSPT-----LSTPSPSLMPTGYKSMELS-----DISNENC 371
* . . *
```

```
Db 416 DSOISTVNVVLKGSPTSPSYAYITGLSNTVSVFNTGNNTLAKTVPGVGNPMGVAISPDT 475
QY 372 RI---NRYGYFRATITIVDGIIL--EVNLIQVADVPIP---TLQPDN-----409
Db 476 RYVYVNTNGY-RGSVSVIDTARGEVITIVDGNKYSPCGIAVTPDGKKLYVSDRDINGV 534
QY 410 SLMDFTVTCGATPTE-----ACTIISDPTQIAONRVCSPPVA-----DELCLLSYRR 458
Db 535 SVIDTSTNTVATVPAGINPLGVAITPDGRKAYVANRYSNVSVIDTVTNNEIAAAVKGT 594
QY 459 -----AFNGSGT--YCVN-----FTLGDDASLALTSAL-----ISIPGKDLGSLRTVN 500
Db 595 GPCGVSFNODGRFLRVNTNCESNSVIDTATVTDTLAVKWPGLCVSPDPGTKLIYAN 654
QY 501 GVLISITGICLAMEVTMTVILLYKKHKTY----KP-----IGNCTRNVVYKGLSVF----546
Db 655 ERSNNVSVIDAETKNTAAIKVGRSPYGIATVDPGTVKVVVANGCNENLGTISIIDTAT 714
QY 547 ---LSHAKAPFS 555
Db 715 NRVIATEKAGFS 726

RESULT 12
Q94C44
ID Q94C44 PRELIMINARY; PRT; 991 AA.
AC Q94C44;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-NAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hydroxyproline-rich glycoprotein VSP4..
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93372571; PubMed=7689982;
RA Waffenschmidt S., Woessner J.P., Beer K., Goodenough U.W.;
RT "Isodityrosine cross-linking mediates insolubilization of cell walls in Chlamydomonas."
RL Plant Cell 5:809-820(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Haas J., Woessner J.P., Waffenschmidt S.;
RT "VSP4-hydroxyproline-rich glycoprotein from Chlamydomonas reinhardtii."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY036106; AAK83527.1; -.
DR InterPro; IPR001412; tRNA-synt.1.
DR PROSITE; PS00178; AA_tRNA_LIGASE_I; UNKNOWN.1.
SQ SEQUENCE 991 AA; 104614 MW; 6B3CB8B3311CD800 CRC64;

Query Match 4.5%; Score 137.5; DB 10; Length 991;
Best Local Similarity 22.8%; Pred. No. 0.01;
Matches 68; Conservative 37; Mismatches 104; Indels 89; Gaps 13;

QY 246 NDRNS-----SDETFRLDLPFIFFDVLHDPHFNLNYSATSYKWNFGDNT-----289
Db 620 NDANSADFSCPSAGADET-ANVTITYDIDCLLVGLHLSGQVYVAMHQITADNNPLYAW 678
QY 290 GLFVSNHNTLNHTYVLNGTFENLTVQTAVPGPCSPTSPSSSTSPSPASSPSPTLS--347
Db 679 GQYQAQNPAGYTPPV-----VATPSPSPSPKPSPPSPSPSPSPSPSPSPS 729
QY 348 ---TPSPSLMPTGYKSMELSDI-----SNENCR-INRYGY----FRATITIVDGIILEVNI 394
Db 730 PSPSPSPSPSPSPSPSPSTNTSQSSTNKGKCVSNWTYLNNOYSDCVTVPGEEKGNW 789
QY 395 IQV-----ADVPITPLQPDNSLMDFTVTCGATPT 424
Db 790 COVDRSNGNCANARNGWWDYCTPSCGTPNAPSPSPSPSPSPSPSTSTGTGTPSCGAGAPT 849
```







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 18:14:00 ; Search time 2476 Seconds  
(without alignments)  
3741.445 Million cell updates/sec

Title: US-09-943-075A-2  
Perfect score: 3061  
Sequence: 1 MESLCGVLLVLLAAGLPQ.....PFSRGDREKDPILLQDPWML 572

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09943075/runat\_07042003\_155631\_17942/app\_query.fasta\_1.711  
-DB=EST -QPMF=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09943075.ecgn\_1.1.1906 -runat\_07042003\_155631\_17942 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2148	70.2	2636	11	BC025297	BC025297 Homo sapi
2	1117	36.5	850	9	AL542812	AL542812 AL542812
3	1045	34.1	810	9	AU139997	AU139997 AU139997
4	972	31.8	880	12	BG769190	BG769190 602743285
5	963	31.5	723	14	BQ694976	BQ694976 1001119 H
6	945	30.9	822	12	BG679094	BG679094 602627047
7	903	29.7	1071	13	BM551384	BM551384 AGENCOURT
8	904	29.5	974	12	BF299741	BF299741 602029248
9	893	29.4	706	14	BM760740	BM760740 K-EST0041
10	896	29.3	704	9	AU135203	AU135203 AU135203
11	895	29.3	943	12	BF139360	BF139360 601785223
12	892.5	29.2	883	9	AL550464	AL550464 AL550464
13	887	29.0	623	10	BB613060	BB613060 BB613060
14	887	29.0	662	10	BB642829	BB642829 BB642829
15	880.5	28.8	790	13	BG966501	BG966501 602833017
16	865.5	28.3	824	9	AU138058	AU138058 AU138058
17	864	28.2	963	12	BG261312	BG261312 602373064
18	859	28.1	804	9	AU140063	AU140063 AU140063
19	858.5	28.0	1022	13	BI155766	BI155766 602904328
c 20	857.5	28.0	732	9	A1640753	A1640753 wa15b03.x
21	849.5	27.8	748	10	AV717139	AV717139 AV717139
22	831	27.1	625	9	AL695277	AL695277 DKFZP313M
23	828	27.0	756	13	BI561155	BI561155 603253684
24	820.5	26.8	818	9	AU135825	AU135825 AU135825
25	818.5	26.7	972	12	BG260561	BG260561 602372115
c 26	812.5	26.5	573	10	BE672596	BE672596 7b7de11.x
27	810	26.5	813	9	AU136173	AU136173 AU136173
28	791	25.8	720	13	BI602015	BI602015 603246239
29	790	25.8	707	13	BI820501	BI820501 603036127
30	788	25.7	718	9	AU136295	AU136295 AU136295
31	784.5	25.6	1107	13	BM550298	BM550298 AGENCOURT
32	782	25.5	495	9	AA510188	AA510188 YG31904.r
33	776	25.4	722	9	AU140013	AU140013 AU140013
34	775.5	25.4	503	10	BB855757	BB855757 BB855757
35	775.5	25.3	1089	13	BM547872	BM547872 AGENCOURT
36	772	25.2	729	12	BF978659	BF978659 602149423
37	763	24.9	670	9	AU136367	AU136367 AU136367
38	762	24.9	737	12	BG769755	BG769755 602744655
39	762	24.9	765	12	BG403585	BG403585 602419294
40	759	24.8	767	9	AU139623	AU139623 AU139623
c 41	753	24.6	583	10	AW368056	AW368056 CM0-HW018
42	753	24.6	959	12	BG288903	BG288903 602383849
43	745.5	24.4	720	12	BF125003	BF125003 601762160
c 44	740.5	24.2	554	9	A1631045	A1631045 tx54h04.x
45	737	24.1	656	9	AU139792	AU139792 AU139792

# ALIGNMENTS

RESULT 1  
BC025297

LOCUS Homo sapiens, glycoprotein (transmembrane) nmb, clone  
DEFINITION IMAGE:4877773, mRNA.

ACCESSION BC025297  
VERSION BC025297.1 GI:19264140  
KEYWORDS HTC.

SOURCE Homo sapiens.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2636)  
AUTHORS Strausberg, R.  
TITLE Direct Submission

## JOURNAL

Submitted (05-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

## REMARK COMMENT

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fiehl, Erin Garland, Ran Guin,

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline

Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 42 Row: 1 Column: 5

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 4505404

This clone has the following problem: frame shifted.

## FEATURES

source

1..2636  
 /organism="Homo sapiens"  
 /db\_xref="LocusID:10457"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:487773"  
 /tissue\_type="Skin, melanotic melanoma, high MDR."  
 /clone\_lib="NIH.MGC.49"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"

BASE COUNT 750 a 586 c 584 g 716 t

## ORIGIN

## Alignment Scores:

Pred. No.: 2,61e-173 Length: 2636  
 Score: 2148.00 Matches: 393  
 Percent Similarity: 82.39% Conservative: 75  
 Best Local Similarity: 69.19% Mismatches: 86  
 Query Match: 70.17% Indels: 14  
 Db: 11 Gaps: 1

US-09-943-075A-2 (1-572) x BC025297 (1-2636)

QY 1 MetGluSerLeuCySGlyValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGln 20  
 Db 52 ATGGAAATGCTCTACTATTTCTCGGGATTTCTCTCTCGCTGGCGAAGATTGCCACTTGAT 111  
 QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
 Db 112 GCCGCCAAACGATTTCATGATGTCGCTGGCAATGAAGACCTCTCTCATAGAGGAG 171  
 QY 41 AsnAsnGlnLeuArgGlyTyrTrpSerSerAspGluAsnGlnTyrAspGluGlnLeuTyrPro 60  
 Db 172 CACAATCAATTAATATGGCTGGCTCTCTGATGAAATGACTTGAATGAATAACTCTTACCCA 231  
 QY 61 ValTrpArgArgGlyGluGlyValGlyTyrLysAspSerTrpGluGlyGlyValGlnAla 80  
 Db 232 GTGTGGAGCGGGAGACATCAGGTGAAATACTCTCGAAGGAGGCGCTGTGCAGCG 291  
 QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
 Db 292 CTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAAATATACATTTGCGGTGACCTG 351  
 QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
 Db 1390 AATGGGCTGGGACGCTACTGTGTGAACCTCACCTGGGGGATGCACACACGCTGGCTCTC 1449

Db 352 ATATTCCCTAGATGCCAAAGAGAGATGCCAATGGCAACATAGTCTATGAGAAGAAGACTGC 411  
 QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140  
 Db 412 AGAAATGAGGCTGGTTTATCTGTCGATCCGTATCTTTACAACATGGACAGCATGGTCAGAG 471  
 QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
 Db 472 GACAGTGGCGGGAATGGCCCGCCCAAGCCCATCAACGCTTTCCTCGATGGGAAA 531  
 QY 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
 Db 532 CCTTTTCTCCACCCCGGATGGAGAGATGGAATTCATCTAGCTCTTCCACACACTT 591  
 QY 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
 Db 592 GGTCAAGTATTCAGAAATGGGACGATGTCAGTGAGAGTTTCTGTGAACACAGCCAAAT 651  
 QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220  
 Db 652 GTGACACTGGGCGCTCACTCATGGAAGTGAATCTCTACAGAGACATGACGGGCATAT 711  
 QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIlePheVal 240  
 Db 712 GTTCCCATCGCAAGTGAAGATGTGTACGTGTAAACAGATCAGATTCTCTGTGTGTG 771  
 QY 241 ThrMetTyrGlnLysAspAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
 Db 772 ACTATGTTCCAGAACAGCATCGAATTCATCCGACCAACCTTCTCAAGATCTCCCC 831  
 QY 261 IlePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
 Db 832 ATATGTTTGTATGTCCTGATTTCATGATCCTAGCCACTTCTCAATATTATTCATCAATAC 891  
 QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
 Db 892 TACAGTGGAGCTTCGGGATTAATCTGCGCTGTGTTTCCACCAATCATACTGTGAAT 951  
 QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
 Db 952 CACAGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAGCTGCAGCACCA 1011  
 QY 321 GlyProCysProSerProSerProSerProSerProSerSerSerSerProSerProAlaSer 340  
 Db 1012 GGACCTTGTCCGCCACCCGCCACCCACCCAGACCT----- 1047  
 QY 341 SerProSerProThrLeuSerThrProSerProSerProSerProThrGlyTyrLysSer 360  
 Db 1048 -----TCAAAACCCACCCCTCTCTTTAGGACCTGCTGTGACAAACCCC 1089  
 QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380  
 Db 1090 CTGGAGCTGAGTAGGATTTCTCTGATGAAATGCGCAGATTAACAGATATGCGCACTTCAA 1149  
 QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400  
 Db 1150 GCCACCATCACAATTTGAGAGGAATCTAGAGGTTAAACATCATCCAGATGACAGACGTC 1209  
 QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420  
 Db 1210 CTGATGCGCGGTGCCATGGCTGAAAGCTCCCTAATAGACTTTCGTGACCTGCCAAGG 1269  
 QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440  
 Db 1270 AGCATTCACCGAGGCTGTGTACCATCATTTCTGACCCCACTCGAGATCACCAGAAC 1329  
 QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460  
 Db 1330 ACAGTCTCAGCCCTGTGGATGTGGATGTGTCTGTCTGCTGACTGTGAGACCACTTC 1389  
 QY 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480  
 Db 1390 AATGGGCTGGGACGCTACTGTGTGAACCTCACCTGGGGGATGCACACACGCTGGCTCTC 1449



Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: [genomics@hri.co.jp](mailto:genomics@hri.co.jp)  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES	source	Location/Qualifiers
	1..810	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="PLACE1011759" /clone_lib="PLACE1" /tissue_type="placenta" /note="Vector: pME18SFL3"
BASE COUNT	229 a 172 g 203 t	5 others

Alignment Scores:			
Pred. No.:	1,66e-79	Length:	810
Score:	1045.00	Matches:	190
Percent Similarity:	60.78%	Conservative:	37
Best Local Similarity:	67.62%	Mismatches:	40
Query Match:	34.14%	Indels:	15
DB:	9	Gaps:	2

US-09-943-075A-2 (1-572) x AU139997 (1-810)

Qy	108	GlulAspAlaAsnGlyAsnIleValTyrGluArgAsnCysArgSerAspLeuGluLeuAla	121
Db	1	AAAGATGCCAATGGCAACATAGTCTATGAGAGAAGCTGCAGAAATAGAGCTGGTTATATCT	60
Qy	128	SerAspProTyrValTyrAsnTrpThrGlyValAspAspGluAspTyrGluAspAsn	147
Db	61	GCTGATCCCATATGTTTACAACTGGACAGCATGTCAGAGACAGTATGGGAAAAATGGC	120
Qy	148	ThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPheProArgProHisGly	167
Db	121	ACGGCGCAAGAGCATCAATACGCTCTCCCTGATGGGAAACCTTTTCCTCACCAACCCGGA	180
Qy	168	ArgLysLysTrpAsnPheValTyrValPheHisThrLeuGlyGlnTyrPheGlnLysLeu	187
Db	181	TGGAGAGAGATGGAAATTCCTACTAGCTCTCCACACACTTGGTCAGTATTTCCAGAAATTG	240
Qy	188	GlyGlnCysSerAlaArgValIserIleAsnThrValAsnLeuThrValGlyProGlnVal	207
Db	241	GGACGATGTTCACTGAGAGAGTTCTGTGAAACACAGCCAAATGTGACACTTTGGCGCTCAACTC	300
Qy	208	MetGluValIleValPheArgArgHisGlyArgAlaTyrIleProIleSerLysValLys	227
Db	301	ATGGAAATGACATGCTCTACAGAAGACATGGAGGGCATATGTTCCCATCGCACAACTGAAA	360
Qy	228	AspValTyrValIleThrAspGlnIleProIlePheValThrMetTyrGlnLysAsnAsp	247
Db	361	GATGTGTACGTGGTAACAGATCAGATCCCTGCTGTTGTGACTATGTTCAGAGAAGACGAT	420
Qy	248	ArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhePheAspValIleIle	267
Db	421	CGAAATTCATCCGAGAAACCTTCTCAAGATGTCCTCCCATATATGTTTGATGTCTCATTT	480
Qy	268	HisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLysTrpAsnPheGlyAsp	287
Db	481	CATGATCCTAGCCACTTCTCTCAATTATTCTACCATTAACCTACAAAGTGAGGCTTCGGGAT	540
Qy	288	AsnThrGlyLeuPheValIserAsnAsnHisThrLeuAsnHisThrTyrValIleAsnGly	307
Db	541	AATACTGGCTGTGTTTTCACCAAAATCATACTGTGAATCACACGATGTGCTCAATGGA	600
Qy	308	ThrPheAsnPheAsnLeuThrValGlnThrAlaValProGlyProCysProSerProThr	327

Db	601	ACCTTACGCTTAACTCACTGTAAGCTGCAGCACCAGGACCTTGTCCG	-----	651
Qy	328	ProSerProSerSerThrSerProSerProAlaSerSerProSerProThrLeuSer	347	
Db	652	-----CCANCGCCA-----	-----	678
Qy	348	ThrProSerProSerLeuMetProThrGlyTyrLySerMetGluLeuSerAspLeuSer	367	
Db	679	AAAGCCCAACCTCTTTAGACCTCTGGGTGA-CAACCCCTGGAGCTGAGTAGATTCTT	737	
Qy	368	AsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrLeuValAsp	387	
Db	738	GATGAACACTGCCAGATTACAGNATATGGCACATTTCAAGCCACCACCATCACAANTTGTANA	797	
Qy	388	Gly 388		
Db	798	GGG 800		
RESULT	4			
LOCUS	BG769190	880 bp	mRNA	linear
DEFINITION	602743285F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4873052 5',			EST 15-MAY-2001
ACCESSION	BG769190			
VERSION	BG769190.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1 (bases 1 to 880)			
JOURNAL	NH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC/DCTD/DP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LUCM1750 row: i column: 21 High quality sequence stop: 792. Location/Qualifiers 1. 880			
FEATURES				
	source			

FEATURES  
SOURCE

```
1. .860
/organism="Homo sapiens"
/db xref="taxon:9606"
```

```

/analyze="IMAGE:4873052"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCRB7; Site_1: xhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACAGAG(C). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH MGC

```

BASE COUNT	255 a	220 c	186 g	219 t
ORIGIN				

Alignment Scores:		
Pred. No.:	3.3e-73	880
Score:	972.00	133
Percent Similarity:	75.97%	41
Best Local Similarity:	82.66%	Conservative:
Query Match:	31.75%	Mismatches: 46
		Indels: 28
		Gaps: 3
		pa. 12

```

US-09-943-075A-2 (1-572) x BG769190 (1-880)

QY 106 GlnLysGluAspAlaAsnGlyAsnIleValThrGluArgAsnCysArgSerAspLeuGlu 125
Db 1 GAAAGAGAGATGCAATGCAATAGTCTATGAGAAAGCAATGAGGCTGGT 60
QY 126 LeuAlaSerAspProTyrValThrAsnTrpThrThrGlyAlaAspAspGluAspTrpGlu 145
Db 61 TTATCTGTCATGTCATATGTTTCACTGAGACAGATGCTCAGAGGACAGTACGGGNA 120
QY 146 AspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPheProArgPro 165
Db 121 AATGACACGGCCGCAAGCCATCAATAGCTCTCCCTGATGGGAACCTTTCTCCACCAC 180
QY 166 HisGly-ArgLysLysTrpAsnThrPheValThrPheHisThrLeuGlyGlnThrPheG1 185
Db 181 CCCGATGAGACAGCATGGAATTTCACTAGCTTCCACACACATTTGGTCAATTTCCA 240
QY 185 nLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsnLeuThrValGlyPr 205
Db 241 GAAATGGGAGATGTTAGTGAGATTTCTGTCAACACAGACCAATGTGACATGTTCCAGAA 420
QY 205 GlnValMetGluValIleValPheArgArgHisGlyArgAlaTyrIleProIleSerLys 225
Db 301 TCAACTCATGGAAGTACTGCTACACAGACATGGACGGCATATGTTCCCATCGCACA 360
QY 225 sValLysAspValTyrValIleThrAspGlnIleProIlePheValThrMetTyrGlnLys 245
Db 361 AGTGAAGATGTTAGCGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAA 420
QY 245 sAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhePheAspVa 265
Db 421 GAACGATCGAATTCATCCGAGCAAACTTCTCAAGATCTCCCATATGTTGATGT 480
QY 265 IleuLeHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLysTrpAsnPh 285
Db 481 CCGATTCATGATGCTAGCCACTTCCCTCAATTTATCTTACCATTAACAGAGTGGAGCTT 540
QY 285 eGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThrTyrValLe 305
Db 541 CGGGATAATAGTGGCTGTTGTTTCCACCAATCATACTGTCAATCACACGATGCTGT 600
QY 305 uAsnGlyThr-PheAsnPheAsnLeuThrValGlnThrAlaValProGlyProCysProS 325
Db 601 CAATGGAACCTTACGCTTAACTCTACTGTGAAAGCTGCAGACACAGGACCTTGTGCGC 660
QY 325 erProThrProSerPro-SerSerSerThrSerProSerProAlaSerSer----- 341
Db 661 CACCTGCACAGCAGCAGAGCTTCAAAACGACCCCTTCTTACCAACTACTCTAAATCT 720
QY 342 -----ProSerProThrLeuSerThrProSerProSerProSerProThrGly 357
Db 721 TATTGATTCAAACATCCCGAGAACTGCTGGTGTGACAAACCCCTGG----- 768
QY 358 TyrLysSerMetGluLeuSerAspIleSerAsnGluAsn-----CysArgIle 373
Db 769 -----GACCTGAGTAAGGATTCCTGTGATGAACACCTTGGCCAGATT 807
QY 374 AsnArg-TyrGlyTyrPheArgAlaThrIle-ThrIleVal-AspGlyIleLeuGluVal 392
Db 808 AACAGATATGGCACTTTCAAGCCAGCATCCCAATTTGTAGCAGCGAATCTTAGAGTCT 867
QY 393 AsnIleIleGln 396
Db 868 AACATCATCCAG 879

RESULT 5
BO694976
LOCUS
DEFINITION 100119 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
CDNA 5', mRNA sequence.
ACCESSION BO694976
VERSION BO694976.1 GI:21820292

KEYWORDS SOURCE
ORGANISM human.
REFERENCE 1 (bases 1 to 723)
AUTHORS Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
TITLE EST analysis of human adipose gene expression
JOURNAL Unpublished (2002)
COMMENT Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR PRIMERS
FORWARD: CTCGGGAAGCGCCGCTTGTGGT
REVERSE: AATGACACTCATAGGGGAATGG
Seq primer: GTTGGTACCGGAATTC.
FEATURES
source
1..723
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/sex="Male and Female"
/tissue_type="Adipose"
/note="vector: lambda triplex"
BASE COUNT 188 a 198 c 157 g 180 t
ORIGIN
Alignment Scores:
Pred. No.: 1,51e-72 Length: 723
Score: 963.00 Matches: 179
Percent Similarity: 81.10% Conservative: 27
Best Local Similarity: 70.47% Mismatches: 33
Query Match: 31.46% Indels: 15
DB: 14 Gaps: 1
US-09-943-075A-2 (1-572) x BO694976 (1-723)
QY 228 AspValTyrValIleThrAspGln-IleProIlePheValThrMetTyrGlnLysAsnAs 247
Db 3 GATCTGATGCTGATCAAGATCAAGATTCCTGCTGTTGTGACTATGTTCAGGAAGACGA 62
QY 247 pAT3AsnSerSerAspGluThrPheLeuArgAspLeuProIlePhePheAspValLeuI 267
Db 63 TCGAATATTCATCCGAGAAACCTTCTCAAGATCTCCCATTTATGTTGCTCTGAT 122
QY 267 eHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLysTrpAsnPheGlyAs 287
Db 123 TCATGATCTGATCCACTTCTCTCAATTTATTCACCATTAACACAGTGGAGCTTCGGGA 182
QY 287 pAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThrTyrValLeuAsnG1 307
Db 183 TAATACTGGCTGTTGTTTCCACCAATCATACTGTGAATCACACGATGTGCTCAATGG 242
QY 307 yThrPheAsnPheAsnLeuThrValGlnThrAlaValProGlyProCysProSerProth 327
Db 243 AACCTTTCAGCTTAACTCTGCTGAAAGCTGCAGACACAGGACCTTGTCCGCCCGCC 302
QY 327 rProSerProSerSerThrSerProSerProAlaSerSerProSerProThrLeuSe 347
Db 303 ACCACCCACACCT-----TC 320
QY 347 rThrProSerProSerLeuMetProThrGlyTyrLysSerMetGluLeuSerAspIleSe 367
Db 321 AAAACCCACCCCTTCTTAGGACCTGCTGTCGTGACAAACCCCTGGAGCTGAGTAGGATTC 380
QY 367 rAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAs 387
Db 381 TGATGAAGAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCAATTTGTA 440

```

```

QY 387 pGlyileLeuGluValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnPr 407
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 441 GGAATCTTTAGAGTTAATCATCATCCAGATGACAGACGCTCCTGATCGCGTGCCATGGCC 500

QY 407 oAspAsnSerLeuMetAspPheIleValThrCysLysGlyAlaThrProThrCiuAlaCy 427
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 501 TGAAGCTCCCTTAATAGACTTTGCGTGACCTGCCAAGGAGGATTCCTCCACGAGGATCTG 560
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 427 sThrIleIleSerAspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAl 447
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 561 TACCATCAATTCGACCCCACTCGAGATCACCCACAGACAGCTCTGCAGCCCTGTGGA 620

QY 447 aValaspGluLeuCysLeuLeuSerValArgArgAlaPheAsnGlySerGlyThrTyrcy 467
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 621 TGTGGATGAATGTCTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTG 680
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 467 sValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 681 TGTGAACCTCACCTCGGGGATGACACAGCCTGGGTTTA 720

RESULT 6
BG679094 822 bp mRNA linear EST 01-MAY-2001
LOCUS 602627047F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4752178 5',
DEFINITION mRNA sequence.
ACCESSION BG679094
VERSION BG679094.1 GI:13910491
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM10609 row: m column: 11
High quality sequence stop: 816.
FEATURES
Location/Qualifiers
1..822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4752178"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 232 a 186 c 204 g 200 t
ORIGIN
Alignment Scores:
Pred. No.: 6,21e-71 Length: 822
Score: 945.00 Matches: 198
Percent Similarity: 83.69% Conservative: 38
Best Local Similarity: 70.21% Mismatches: 32
Query Match: 20.87% Indels: 15
DB: 12 Gaps: 0
US-09-943-075A-2 (1-572) x BG679094 (1-822)
QY 22 AlaLysArgPheAspValLeuGlyHisGluGlnTyrProAspHisMetArgGluAsn 41

```

```

Db 2 GCCAAACGATTTTCATGATGCTGGGCAATGAAGACCTTCTGCTTACATGAGGAGCAC 61
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 42 AsnGlnLeuArgGlyTrpSerSerGluAsnGluTrpAspGluGlnLeuTyrProVal 61
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 AATCAATTAATAGGCGGCTTCTGATGAATGACTGGAATGAAACACTCTACCCAGTG 121
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 TrpArgArgGlyGluGlyArgTrpLys-AspSerTrpGluGlyArgValGlnAlaAl 81
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 TGAAGCGGGAGACATGAGGTGGAACAACCTCTCTGGAAGGAGCGCTGTGCGAGGGGT 181
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 81 alaThrSerAspSerProAlaLeuValCysSerAsnIleThrPheValValAsnLeuVa 101
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 CCTGACCACTGACTACCAAGCCCTCGTGGCTCAAAATATAACATTCGGGTGAACCTGAT 241
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 101 lPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCysAr 121
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 ATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTCGAG 301
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 q-SerAspLeuGluLeuAlaSerAspProTyrValTyr-AsnTTPThrThrGlyAlaAsp 140
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 AATATGAGCGCTGGTTTATCTGCTGATCCGTATGTTTACAAGCTGGACAGCATGGTTCAGAG 361
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 141 AspGluAspTrpGluAspAsnThrSerGln-GlyGlnHisLeuArg-PheProAsp-Gly 159
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 GACAGTGACGGGGAATAATGGCACCGGCCCAAGCCCTATCATAAAGCGCTCTCCCTGATGGT 421
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 160 LysProPheProArgProHisGlyArgLysLys-TripAsnPhe-ValTyrValPheHis 178
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 422 AAACCTTTTCCCTCACCCACCCCGGATGGCAGAGATGGAACTTCAGTCTAGCTTCCAC 481
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 179 ThrLeuGlyGlnTyrPheGlnLys-LeuGlyGlnCysSerAlaArgValSerIleAsnTh 198
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 ACATTGGTCAGTATTTCAGAAAGATTGGGACGATGTTCACTGAGAGTTCCTGTGGAACAG 541
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 198 r-ValAsnLeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyA 218
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 CAGCCATGTGACACTTGGGCTCAACTCATGGAAGTACTGTCTACAGAGACATGGAC 601
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 218 rGalaTyrIleProIleSerLysVal-LysAspValTyrValIleThrAspGlnIlePro 237
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 602 GGGCATATGTCCCATCGCACAAAGTGACAAGATGTGTACGTGGTAACAGATCAGATTCTC 661
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 lIlePhe-ValThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuAr 257
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 662 GTGTACTGTGACTATGTTCCAGAAAGACATCGAATTCATCCGACGAACCTTCCTCAA 721
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 257 gAspLeuProIlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSe 277
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 722 AGATCTCCCATTTATGTTTGTGATGCTCTGATTCATGATCTAGGCACCTTCCTCAATTATTC 781
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 277 rAlaIleSerTyrLysTrpAsnPhe-GlyAspAsnThr 289
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 782 TACCATTAC-TACAAGTGGAGCTTCGGGGGATATAACT 818
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 7
BM551384
LOCUS BM551384.1 GI:18788394
DEFINITION 5' mRNA sequence.
ACCESSION BM551384
VERSION BM551384.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1071)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

```

Tissue Procurement: ATCC/DCTD/DTP  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1AM12256 row: a column: 16  
 High quality sequence stop: 532.  
 Location/Qualifiers  
 1..1071  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH\_MGC\_72"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."  
 BASE COUNT 288 a 256 c 269 g 258 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,06e-67 Length: 1071  
 Score: 909.00 Matches: 173  
 Percent Similarity: 81.15% Conservative: 38  
 Best Local Similarity: 66.54% Mismatches: 44  
 Query Match: 29.70% Indels: 5  
 DB: 13 Gaps: 1

US-09-943-075A-2 (1-572) x BM551384 (1-1071)

Qy 1 MetGluSerLeuGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20  
 Db 102 ATGAATGCTCTACTACTATTCCTGGGATTCCTGCTGGCTGGCAAGATGCCACTTGAT 161  
 Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
 Db 162 GCCGCCAAAGCATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 221  
 Qy 41 AsnAsnGlnLeuArgGlyTyrPdsSerAspGluAsnGluTyrAspGluLeuTyrPro 60  
 Db 222 CACATCAATTAATGCTGCTCTCTGATGAAATGACTGGAATGAAACTTACCCCA 281  
 Qy 61 ValTrpArgGlyGlyGlyArgTyrLysAspSerTrpGluGlyGlyValGlnAla 80  
 Db 282 GTGTGGAAGCGGGAGACATGAGGTGAAAACTCTCTGGAAGGAGGCGGTGTCAGGCG 341  
 Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
 Db 342 GTCCTGACCACTGACTCACCAGCCCTGCTGGGCTCAATATTAACATTTGCGGTGAACCTG 401  
 Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
 Db 402 ATATTCCCTAGATGCCAAAGGAAGATGCAATGCGCAACATCTCTATGAGAAGAACTGC 461  
 Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAsp 140  
 Db 462 AGAATGAGGCTGTTTATCTGCTGATCCATATGTTTAACTGACGACATGTCAGAG 521  
 Qy 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
 Db 522 GACAGTGACGGGAAATGACCCGCCAAGCCATCATACGCTCTCCCTGATGGGAA 581  
 Qy 161 ProPheProArgProHisGlyArgLysLysTyrAsnPheValTyrValPheHisThrLeu 180  
 Db 582 CTTTTCTCACCAACCCCGATGGGAAGATGGAATTTTCATCTAGCTCTCCACACACTT 641  
 Qy 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
 Db 642 GGTCAGTATTTCCAGAAATTTGGGACCATGTTTCAGTCAGAGAGTTTCTGTGAACACAGCAAT 701

Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArg-AlaTy 220  
 Db 702 GTACACATTTGGCGCTCAACTCATGAGTGCCTGTACAGAAAGACATGGGAGGCGCAT 761  
 Qy 220 rLeProileSerLysVal-LysAspValTyrVal-IleThrAspGlnIleProIlePhe 239  
 Db 762 TGTTCCTCATCGCACAAAGTGGAAAGATGTGTACGTGGGAAACACATCAAAATCCCTGTGT 821  
 Qy 240 ValThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArg 257  
 Db 822 TTGGGACTATGTTCCCGAAGGACCGA-----TCCGAAATTTCTTCCGGAAGA 869  
 RESULT 8  
 BF299741  
 LOCUS  
 DEFINITION EF299741 974 bp mRNA linear EST 21-NOV-2000  
 G02029248F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4164706 5',  
 mRNA sequence.  
 ACCESSION BF299741  
 VERSION BF299741.1 GI:11246264  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 REFERENCE 1 (bases 1 to 974)  
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1AM9450 row: o column: 11  
 High quality sequence stop: 665.  
 Location/Qualifiers  
 1..974  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_lib="NCI\_CGAP\_SG2"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 252 a 231 c 289 g 202 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.51e-67 Length: 974  
 Score: 904.00 Matches: 189  
 Percent Similarity: 76.75% Conservative: 19  
 Best Local Similarity: 69.74% Mismatches: 46  
 Query Match: 29.53% Indels: 17  
 DB: 12 Gaps: 4

US-09-943-075A-2 (1-572) x BF299741 (1-974)

Qy 1 MetGluSerLeuGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20  
 Db 79 ATGGAAGCTCTCTGGGGTCTGGGATTTCTGCTGGCTGCGAGACTGCTCTCCAG 138  
 Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
 Db 139 GCTG-CAAGCCGATTCGTGTGATGTCTGGGCGCATGAACAGTATCCCAATCACATGAGAG 198

Qy	41	AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGlnLeuTyrrPro	60
Db	199	CACACCAATATACGGTGGTCTTCGGATGAATAATGGATGGATGAACACCTGTATCCA	258
Qy	61	ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla	80
Db	259	GTCTGGAGGAGGGACACGGCAGTGGAAAGACCTCTGGCAAGAGAGCCCTGCAGCCA	318
Qy	81	AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
Db	319	GTCTTCACCAAGTACCTACCGGCTCTGGTGGGTTCCAATATCACCTTTGTGTGAACCTG	378
Qy	101	ValPheProArgCysGlnLysGlnAlaAsnGlyAsnIleValTyrgluArgAsnCys	120
Db	379	GTCTTCCCAGATGCCAAGAAGAGATGCTAATGGCAATATGCTCTATCAGAAGAAGTGC	438
Qy	121	ArgSerAspLeuGlnLeuAlaSerAspProTyrrValTyrrAsnTrpThrThrGlyAlaasp	140
Db	439	AGGAATGATTTGGGACTGACCTTGACCTGCATGCTCTACAACTGGACTCGAGGGCGAGAT	498
Qy	141	AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAsp-Glyly	160
Db	499	GATGGTGACTGGGAAGATGGACACCGCAAGCCAGCATCTCAGGTTCCTCCGGCAAGGAG	558
Qy	160	sPropPheProArgProHis-GlyArgLysLysLysTrpAsnPheValTyrrVal-PheHisThr	179
Db	559	GCCTTCCCTCGCCCACTGATGTAAGCAAAATGGACTGGTCTACGTCTTTCACACACA	618
Qy	180	LeuGlyGlnTyrrPheGlnLysLeuGlyGln---CysSerAlaArgValSerLeuAsnThr	198
Db	619	ACTGGCGAGTATTTCCAAAACACTGGTGGTGGTTCACGACACGGGTTTTCTCAATAAACA	678
Qy	199	ValAsnLeuThrValGlyProGlnValMetGluValIle-----ValPheArg	214
Db	679	GCAGGTCCAACTGGGACACGTGGACCTCAGTCCCTGGGAACGTGGACTGCTTCGG	738
Qy	215	-ArgHisGlyArgAlaTyrrIle-ProIleSerLysValLysAspValTyrrValIleThrA	234
Db	739	AAGCATAGACGGGGATTCATTCCTCCAAATAGAAGGGTACAGATGTGTTGGGATACAC	798
Qy	234	sp---GlnIleProIlePheValThrMetTyrrGlnLysAsnAsp-----	247
Db	799	AGATCAAAATCCGTTATTCTGTCGACCATGCCCCCAAGGACTGACACGGGACCTGTTCGA	858
Qy	248	--ArgAsnSerSerAspGluThr	254
Db	859	TAAGAAGATCTCAAGAACAACACA	881
RESULT	9		
LOCUS	BM760740	706 bp mRNA linear EST 04-MAR-2002	
DEFINITION	K-EST0041401 S12SNU216 Homo sapiens CDNA clone S12SNU216-22-G02 5'		
ACCESSION	BM760740	mRNA sequence.	
KEYWORDS	BM760740.1 GI:19090355		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 706)		
JOURNAL	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.		
COMMENT	21C Frontier Korean EST Project 2001		
	Unpublished (2002)		
	Contact: Kim YS		
	Genome Research Center		
	Korea Research Institute of Bioscience & Biotechnology		
	52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea		
	Tel: +82-42-860-4470		
	Fax: +82-42-860-4409		
	Email: yongsung@mail.kribb.re.kr		





```

FEATURES
Source
Location/Qualifiers
1..943
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4013356"
/clone_lib="NCI-CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/Note="Organ: lung; Vector: pCMV-Sport6; Site: 1: NotI;
Site 2: SalI; transgenic model WNT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 221 a 227 c 290 g 205 t
ORIGIN
Alignment Scores:
Pred. No.: 1.16e-66 Length: 943
Score: 896.00 Matches: 186
Percent Similarity: 86.44% Conservative: 18
Best Local Similarity: 78.81% Mismatches: 27
Query Match: 29.27% Indels: 9
DB: 12 Gaps: 2
US-09-943-075A-2 (1-572) x BF193360 (1-943)
Qy 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGln 20
Db 37 ATGGAAAGTCTCTGGGGGCTCTGGGATTCCTGCTGCTGGCTGGAGCTGCTCTCCAG 96
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
Db 97 GCTGCCAAGCGATTTCGTGATGCTGCTGGCCCAAGCAAGATGATCTGATCATCATGAGAG 156
Qy 41 AsnAsnGlnLeuArgGlyTyrSerSerAspGluAsnGlnTyrPaspGluGlnLeuTyrPro 60
Db 157 CACAACCAATTACGTGGCTGCTGCGATGAAATGAAATGAAATGAAATGAAATGAAATG 216
Qy 61 ValTrpArgGlyGlyGlyValGlyTyrPaspGlyTyrPaspGlyTyrPaspGlyTyrPasp 80
Db 217 GTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 276
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 277 GTCTGACACAGTACTACCGCCCTGCTGGGTTCCATATATCATCTTGTGGTGAACCTG 336
Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120
Db 337 GTGTTCCCGAGATGCCAAGGAGAGATGCCAATGCAATATCGTCTATGAGAAGAACTGC 396
Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrThrGlyAlaAsp 140
Db 397 AGGAATGATTG-GGACTGACTTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
Qy 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
Db 456 GATGGTGACTGGGAAGATGGACCAAGCCAGCCAGCATCTCAGATTCCTCCGGACAGGAG 515
Qy 161 ProPheProArgProHisGlyValGlyTyrPaspPheValPheHisHisHisHisHis 180
Db 516 CCCTTCCTCCCGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 574
Qy 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Db 575 GGTCAGTATTC-CAAAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 633
Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArg---ArgHisGly-ArgAl 219
Db 634 TTACAGCTGGG-CCACAGCTATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
Qy 219 atyrilleProfileSerLysValLysAspValTyrValIleThrAsp 234
DB: 12 Gaps: 2
US-09-943-075A-2 (1-572) x AL550464 (1-883)
Qy 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGln 20
Db 65 ATGGAATGCTCTTACCTTCTGCTGGGATTCCTGCTGCTGGCTGCTGCTGCTGCTGCTGCT 124
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
Db 125 GCCGCCAAACGATTCATGATGCTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 184
Qy 41 AsnAsnGlnLeuArgGlyTyrPaspSerAspGluAsnGluTyrPaspGluGlnLeuTyrPro 60
Db 185 CACAATCAATTAATCGTGGTCTTCTGATGAAATGACTGGAATGAAATGAAATGAAATGAA 244
Qy 61 ValTrpArgGlyGlyGlyArgTyrPaspSerTrpGlyGlyGlyGlyGlyGlyGlyGlyGly 80
Db 245 GTGTGGAAGCGGGGAGACATGAGTGGAAAAAAGTCTCTGGAAGGAGGCGGTGTGCGAG 304
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 305 GTCTTGACCAAGTACCTACCAAGCCCTCGTGGGCTCAATATATATATATATATATATAT 364
Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120

```

```

Db 693 CCATCTC-----GAAGTGAACGAAGTGTCTGTGATTACAGAT 729
RESULT 12
AL550464 883 bp mRNA linear EST 16-FEB-2001
LOCUS AL550464 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI057YD14 5
DEFINITION prime, mRNA sequence.
ACCESSION AL550464
VERSION AL550464.1 GI:12887464
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 883)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
E-mail: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..883
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/Note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 243 a 192 c 218 g 219 t 11 others
ORIGIN
Alignment Scores:
Pred. No.: 2.13e-66 Length: 883
Score: 892.50 Matches: 175
Percent Similarity: 79.62% Conservative: 36
Best Local Similarity: 66.04% Mismatches: 52
Query Match: 29.16% Indels: 6
DB: 9 Gaps: 2
US-09-943-075A-2 (1-572) x AL550464 (1-883)
Qy 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGln 20
Db 65 ATGGAATGCTCTTACCTTCTGCTGGGATTCCTGCTGCTGGCTGCTGCTGCTGCTGCTGCT 124
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
Db 125 GCCGCCAAACGATTCATGATGCTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 184
Qy 41 AsnAsnGlnLeuArgGlyTyrPaspSerAspGluAsnGluTyrPaspGluGlnLeuTyrPro 60
Db 185 CACAATCAATTAATCGTGGTCTTCTGATGAAATGACTGGAATGAAATGAAATGAAATGAA 244
Qy 61 ValTrpArgGlyGlyGlyArgTyrPaspSerTrpGlyGlyGlyGlyGlyGlyGlyGlyGly 80
Db 245 GTGTGGAAGCGGGGAGACATGAGTGGAAAAAAGTCTCTGGAAGGAGGCGGTGTGCGAG 304
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 305 GTCTTGACCAAGTACCTACCAAGCCCTCGTGGGCTCAATATATATATATATATATATAT 364
Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120

```









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 10:06:23 ; Search time 4818 Seconds

(without alignments)

3455.128 Million cell updates/sec

Title: US-09-943-075a-2

Perfect score: 3061

Sequence: 1 MESLCGVLFVLLAAGLPQ.....PFSRGDRKDPDLLQKPMWL 572

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh

-Q/cgn2\_1/USPTO.spool/US09943075/runat\_07042003\_155630\_17927/app\_query.fasta\_1.711

-DB=GenEmbl -OPWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US09943075 -CGN\_1\_1\_2586 -runat\_07042003\_155630\_17927 -NCPU=6 -ICPU=3

-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120

-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pi:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_ats:\*  
12: gb\_av:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_pa:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_ov:\*  
22: em\_lov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_ats:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3061	100.0	2320	10	AF184983 Rattus no
2	2705	88.4	2305	10	MMU251685 Mus muscu
3	2703	88.2	2213	6	AR156839 Sequence
4	2703	88.2	2279	10	AF322054 Mus muscu
5	2703	88.2	2299	10	BC026375 Mus muscu
6	2163	70.7	2662	9	AF322909 Homo sapi
7	2163	70.7	2669	6	AR018808 Sequence
8	2163	70.7	2669	6	AR018814 Sequence
9	2163	70.7	2669	9	HSNNB X76534 H. sapiens N
10	2162	70.6	2683	6	AX358788 Sequence
11	2162	70.6	2693	6	AX352281 Sequence
12	2157	70.5	2787	9	BC032783 Homo sapi
13	1481.5	48.4	2467	5	CJQNR71 X94144 C. japonica
14	941	30.7	620	6	AR018825 Sequence
15	870	28.4	162541	2	AC098114 Rattus no
16	823	26.9	104229	2	AC129735 Sequence
17	816	26.7	190519	2	AC110271 Rattus no
18	729	23.8	524	6	AR018817 Sequence
19	719.5	23.5	1690	9	BC011595 Homo sapi
20	717	23.4	473	6	AR018827 Sequence
21	640.5	20.9	2470	5	D88348 Chicken mRN
22	629	20.5	2026	9	HSU01874 Human me20m
23	627.5	20.5	2114	9	HUMGPM5 M32295 Human 95 kD
24	627.5	20.5	2115	6	A45993 Sequence 1
25	627.5	20.5	2130	6	AR167365 Sequence
26	627.5	20.5	2130	6	AX274950 Sequence
27	627.5	20.5	2130	6	AX354933 Sequence
28	627.5	20.5	2130	9	S73003 gp100-melan
29	627.5	20.5	2134	9	BC001414 Homo sapi
30	627.5	20.5	2758	9	AK092881 Homo sapi
31	627	20.5	2172	6	AR063067 Sequence
32	627	20.5	2172	6	AR091800 Sequence
33	627	20.5	2172	6	AR162997 Sequence
34	622.5	20.3	1986	6	AX133406 Sequence
35	622.5	20.3	1986	6	AX133655 Sequence
36	622.5	20.3	1986	6	AX192347 Sequence
37	622.5	20.3	2534	6	AX133528 Sequence
38	622	20.3	2131	6	AX474662 Sequence
39	622	20.3	2131	9	HUMPMEL M77348 Human Pmel
40	620	20.3	1881	6	AX474660 Sequence
41	620	20.3	1881	10	MMU14133 Mus musculu
42	613.5	20.0	158105	2	AC023375 Homo sapi
43	613.5	20.0	169739	9	AC005082 Homo sapi
44	524	17.1	335	6	AR018809 Sequence
45	524	17.1	335	6	AR018811 Sequence

ALIGNMENTS

RESULT 1





```

|||||
Db 1255 GCCACCATCAAAATTTAGATGGAATCCTAGAAAGTCAACATCATCCAGGTAGCAGATGTC 1314
QY 401 ProfileProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
|||||
Db 1315 CCAATCCCACTACAGCTGACAGCTGACAACTACACTGATGGAGTTCATTGTGACCTGCAAGGG 1374
QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440
|||||
Db 1375 GCCACTCCACAGGAACCTGTACGATCATCTCTGACCCACCATGCCAGATGCCGCCAAGC 1434
QY 441 ArgValCysSerProValAlaValAlaAspGluLeuCysLeuLeuSerValArgArgAlaPhe 460
|||||
Db 1435 AGGTGTGACGCCCGGTGCTGTGATGAGCTGTCCCTCTGCTGAGAGAGCCTTC 1494
QY 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
|||||
Db 1495 AATGGGTCCGGCACGTACTGTGTGAATTTCACTCTGGCAGACGATGCAAGCCTGGCCCTC 1554
QY 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500
|||||
Db 1555 ACCAGGCCCTGATCTCTATCCCTGGCAAGACCTAGGCTCCCTCTGAGAACAGTGAAT 1614
QY 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520
|||||
Db 1615 GGTGTCTGATCTCCATTGGCTGCCTGGCCATGTTTGTCCACCATGGTTACCATCTTCTCTG 1674
QY 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540
|||||
Db 1675 TACAAAACACAAAGACGTACAAAGCAATAGGAACTGCACACAGAAACGTGGTCAAGGGC 1734
QY 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560
|||||
Db 1735 AAAGGCTGAGTGTCTTCTCAGCCATGCAAAAGCCCGTCTCCCGAGGAGACCGGGAG 1794
QY 561 LysAspProLeuGlnAspLysProThrMetLeu 572
|||||
Db 1795 AAGATCCACTGCTCAGGACCAAGCATGGATGCTC 1830

RESULT 2
MMU251685
LOCUS
DEFINITION
Mus musculus mRNA for putative transmembrane glycoprotein (NM_
gene).
ACCESSION
AJ251685
VERSION
AJ251685.1 GI:6688787
KEYWORDS
NMB gene; transmembrane glycoprotein.
SOURCE
Mus musculus.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 2305)
AUTHORS
Bachner,D., Schroder,D. and Gross,G.
TITLE
Isolation and developmental expression analysis of murine nmb
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2305)
AUTHORS
Bachner,D.
TITLE
Direct Submission
SUBMITTED (25-NOV-1999)
Bachner D., Zellbiochemie und klinische
Neurobiologie, Universitaetsklinik Eppendorf, Martinistr. 52, 20246
Hamburg, GERMANY
FEATURES
Location/Qualifiers
1..2305
/organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="embryonic"
91..1815
/gene="NMB"
91..1815
/gene="NMB"
/function="unknown"
/codon_start=1
/evidence=experimental
source

```

```

/product="putative transmembrane glycoprotein"
/protein_id="CA665272.1"
/db_xref="GI:6688788"
/translation="MESLCGVGLFLLAAGLPLOAKRFRDVLGHEQYPOHREHNL
RVSSEDENEWDEHLYPVMRGGRWRKSDNEGRVQAVLTSDPALGVNSITFVYNLVF
PROCKEDANGNIYVYKNCNDLGLTSDLVYVNTWAGDDGDMEDGTSRQHLRPDDR
PPRPKRWKWSFVYVHTLQGYFQKLGRCASARVINTVNLTAGQVMEYTVFRYGR
AYTPIKSKVDVYITDQIPVFTVMSQKNDRLNLSDEIFLRLDLPVFDVLIHDPHFLND
SALYSKWNFGDNTGLFVSNHTLNHTYVNLNGTENLNTVQTPVGPCPPSPSPSPSP
STPLDPSPLPTLTPSPSLMPTGYKSMELSDLSNENCRINRYGYFRVAVITVEGTL
EVSIMQIADVPMPPTPOPANSLMDFTYKCGATPMENCTIISDPKCOIAQNRVCSVAV
DGCLLSVRAEFGSGTCVNTLGGDASLALTSTLISPKGDPSPRAVNGVLIISI
GCUAVLYVTWITILLYKKHAYKPIGNCPRNTVKGUSVLLSHAKAPFFRGDQEKDPL
LQDKPRTL"
BASE COUNT 558 a 612 c 584 g 551 t
ORIGIN
Alignment Scores:
Pred. No.: 2,55e-175 Length: 2305
Score: 2706.00 Matches: 508
Percent Similarity: 92.47% Conservative: 20
Best Local Similarity: 88.97% Mismatches: 41
Query Match: 88.40% Indels: 2
DB: 1 Gaps: 1
US-09-943-075A-2 (1-572) x MMU251685 (1-2305)
QY 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20
Db 91 ATGAAAGTCTCTCGGGGGTCTGGGATTTCTGCTGCTGCGAGGACTGCTCCAG 150
QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
Db 151 GCTCCCAAGCATTTCTGATGCTGGGCGCATGACAGTATCCCGATCACATGAGAGAG 210
QY 41 AsnIleGlnLeuArgGlyTyrSerSerAspGluAsnGluTyrAspGluGlnLeuTyrPro 60
Db 211 CACAACCAATTTACGTGGTCTGCTCGGATGAAATGAATGAGATGAACACCTGTATCCA 270
QY 61 ValTyrArgArgGlyGluGlyValArgTyrLysAspSerTyrGluGlyArgValGlnAla 80
Db 271 GTGTGGAGGAGGAGAGAGCGGCGGAGGACTCTCGGAGGAGGCGCTGTGGCAGGCA 330
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 331 GTCCTGACCACTGACTCACCAGGCTCTGGTGGTTCCTCAATATCATCTTTGTGTGAACCTG 390
QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyValTyrGluArgAsnCys 120
Db 391 GTGTTCCTCCAGATCCAGAGAGAGATGCTATGCGCAATATCTCTATGAGAGAACTGC 450
QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnThrThrThrGlyAlaAsp 140
Db 451 AGGAATGATTTGGGACTGACATCTGACCTGCATGCTTACAACCTGGACTCGAGGGCAGAT 510
QY 141 AspGluAspTyrGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
Db 511 GATGCTGACTGGGAGATGGCAGCCGCCGACATCTCAGGTTCCCGGAGAGGAG 570
QY 161 ProPheProArgProHisGlyArgLysLysTyrAsnPheValTyrValPheHisThrLeu 180
Db 571 CCCTTCCTCGCCCCCATGGATGGAAGAAATGAGCTTTGTCTACGCTTTTCACACACTT 630
QY 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Db 631 GCCCAGATTTTCCAAAACCTGGGTGCTGTCAGCAGCGGTTCCTATATAACACAGCTCAAC 690
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220
Db 691 TTGACAGCTGGCCCTCAGGTGCTGGAAGTCACTGCTCTTTCGAAGATACGCGCGGCATAC 750
QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240
|||||

```

Db 751 ATTCCCATCTCGAAGGTGAAGATGCTATGTGATACAGATCAGATCCCTGATTCGTG 810  
Qy 241 ThrMetTyrGlnLysAsnAspArgAsnSerAspGluThrPheLeuArgAspLeuPro 260  
Db 811 ACCATGTCAGAGAAAGATGACAGAACTTCTCTGATGAGATCTTCTCAGACACCTCC 870  
Qy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
Db 871 ATGCTTTCGATGTCCTCATTCATGATCCAGGCACCTTCTCAACGACCTGCCAATTC 930  
Qy 281 TyrLysTyrAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAspHisThrLeuAsp 300  
Db 931 TACAAGTGAACCTTGGGGAACACACCTGCTGTGTCTCCAAACATCAGACCTTGAAT 990  
Qy 301 HisThrTyrValLeuAsnGlyThrPheAspPheAsnLeuThrValGlnThrAlaValPro 320  
Db 991 CACACTTATGTGCTCAATGGAACCTTCACCTTAACCTTACCTGCAACCTGACGTGCC 1050  
Qy 321 GlyProCysProSerProThrProSerProSerSerSerSerSerSerProSerPro 338  
Db 1051 GGGCCATGCCCTCCCTTCGCCCTTCGACCTCCGCTTCACCTTCAACTCCGCCCTTACCT 1110  
Qy 339 AlaSerSerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 358  
Db 1111 TCGCCCTCACCTTTCGCCACATTCATCAACACCTAGCCCTCTTAAATGCTACTGCTTAC 1170  
Qy 359 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 378  
Db 1171 AATCCATGGAGCTGAGTGACATTTCCAATGAAACTGCGGAATAAACAGATATGCTAC 1230  
Qy 379 PheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAla 398  
Db 1231 TTCAGAGCCACCATCACAATTTGTAGAGGGATCTCGAAGTCAGCATCATGAGATAGCA 1290  
Qy 399 AspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCys 418  
Db 1291 GATGTCCTCCATGCCACACCCGCTCCCACTCCCTGATGACATCTACTGTGACCTGC 1350  
Qy 419 LysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 438  
Db 1351 AAAGGGGCGACCCCGATCGAAGCGCTGACGATCATCTCCGACCCCACTGCCAGATCGCC 1410  
Qy 439 GlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArg 458  
Db 1411 CAGAACCGGGTCTGCAGCCCTGTGGCTGTGGATGGCTGTGCTGTCTGTGAGAGA 1470  
Qy 459 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeu 478  
Db 1471 GCCTTCAATGGGTCTGGCACCCTACTGTGTGAATTTCACTCTGGGAGATGATCAAGCC 1530  
Qy 479 AlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThr 498  
Db 1531 GCCCTCACCAGCACCTGATCTATCCCTGGCAAGACCCAGACCTCCCTCTGAGAGCA 1590  
Qy 499 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIle 518  
Db 1591 GTGAATGTGTCTGATCTCCATCGGTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1650  
Qy 519 LeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValVal 538  
Db 1651 TTGCTGTACAAAAACACAGGCGCTACAGCCCAATAGGAACCTGCCCGAGACAGGCTC 1710  
Qy 539 LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAsp 558  
Db 1711 AAGGGCAAGGGCTGAGTGTCTCTCAGTCAGCGCAAGACCCCGTCTCTCCGAGGAGAC 1770  
Qy 559 ArgGluLysAspProLeuLeuGlnAspLysPro 569  
Db 1771 CAGGAGAAGGATCCATTCCTCCAGGACAAGCCA 1803  
RESULT 3  
ARI56839  
LOCUS AR156839 2213 bp DNA linear PAT 08-AUG-2001

DEFINITION Sequence 27 from patent US 6242419.  
ACCESSION AR156839  
VERSION AR156839.1 GI:15125543  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2213)  
AUTHORS Strachan, L., Sleeman, M., Abernethy, N., Onrust, R., Kumble, A. and  
Murison, G.  
TITLE Compositions isolated from stromal cells and methods for their use  
JOURNAL Patent: US 6242419-A 27 05-JUN-2001.  
FEATURES  
Location/Qualifiers  
source 1..2213  
BASE COUNT 526 a 599 c 555 g 533 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.2e-175 Length: 2213  
Score: 2700.00 Matches: 507  
Percent Similarity: 92.29% Conservative: 20  
Best Local Similarity: 88.79% Mismatches: 42  
Query Match: 88.21% Indels: 2  
DB: Gaps: 1  
US-09-943-075A-2 (1-572) x AR156839 (1-2213)  
Qy 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20  
Db 23 ATGGAAGTCTCTGGGGGCTCTGGGATTTCTGCTGCTGGCTGGAGGACCTGCTCTCCAG 82  
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
Db 83 GCTGCCAAGCGATTCGTGATGCTGGGCCATGAACAGTATCCCAATCATCATGAGAG 142  
Qy 41 AsnAsnGlnLeuArgGlyTyrPheSerSerAspGluAsnGluTyrAspGluGlnLeuTyrPro 60  
Db 143 CACAACCAATTCGCTGGCTGCTGCGATGAAATGAATGGATGGATGAACACCTGTATCCA 202  
Qy 61 ValThrArgGlyGlyGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla 80  
Db 203 GTGTGGAGGAGGGAGAGCGCAGCTGGAAGGACTCTCTGGGAAGGAGGCGGTGTGAGGCA 262  
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
Db 263 GTCTGACCACTGACTCACCCTGCTGCTGGGTCCATATACCTTTGTGTGAACCTG 322  
Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
Db 323 GTCTTCCCGAGATGCCAGAAAGAGATGCTATGCAATATCGTCTATGAGAAGAACTGC 382  
Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrGlyAlaAsp 140  
Db 383 AGGAATGATTTGGGACTGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442  
Qy 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
Db 443 GATGGTCACTGGGAAGATGGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 502  
Qy 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
Db 503 CCCTTCCCTCGCCCGCATGGATGGAAGAAATGGAGCTTTGCTACGCTTTTACACACACTT 562  
Qy 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
Db 563 GGCCAGTATTTCCAAAACCTGGGTGCTTCCAGCAGCGGTTTCTATAAACACAGACTCAAC 622  
Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220  
Db 623 TTGACAGCTGGCCCTCAGGTCATGGAAGTACTGCTCTTTCGAAGATACGCCCGCGGATAC 682  
Qy 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240

```

Db 683 ATTCCCATCTCGAAGGTGAAGATGTGTATGTGATAACAGATCAGATCCCTGATTCTG 742
Qy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
Db 743 ACCATGTCCTCCAGAAAGATGACAGAACTTGTCTGATGAGATCTCTCTCAGAGACCTCC 802
Qy 261 IlePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
Db 803 ATCGTCTTCGATGTCCTCATTCATGATCCAGCCACTTCTCAACGACTCTGCCATTGCC 862
Qy 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
Db 863 TACAAGTGGAACTTTGGGGCAACACTGGCTGTGTCTCTCCAAATCACACTTTGAAT 922
Qy 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnThrValGlnThrAlaValPro 320
Db 923 CACACTTATGTGCTCAATGAACTTCAACCTTAACTTACCGTGCAGAACTGCAGTGC 982
Qy 321 GlyProCysProSerProThrProSer-----ProSerSerSerThrSerProSerPro 338
Db 983 GGGCATGCGCTCCCGCTTCCCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 1042
Qy 339 AlaSerSerProSerProThrLeuSerThrProSerProSerSerLeuMetProThrGlyTyr 358
Db 1043 CCGCCCTCACCTCTGCCCACTTATCAACACCTTAGCCCTCTTTAATGCTACTGCTTAC 1102
Qy 359 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 378
Db 1103 AAATCCATGAGCTGAGTACATTTCAATGAATAAAGTCCGGAATAAAGATATGCTTAC 1162
Qy 379 PheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAla 398
Db 1163 TTACAGCCACCATCAATTTAGAGGGGATCTCGAAGTCAGCATCATGCAGATAGCA 1222
Qy 399 AspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCys 418
Db 1223 GATGTCCTCCATGCCACCGCAGCCTGCGCAACTCCCTGATGACTTCACTGTGACCTGC 1282
Qy 419 LysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 438
Db 1283 AAGGGGCCACCCCATGGAAGCTCTAGATCATCTCCGACCCACCTGCCAGATCGCC 1342
Qy 439 GlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArg 458
Db 1343 CAGAACCGGTCTGCAAGCCCTGTGGCTGTGGATGGCTGTGCTGTGCTGTGAGAGA 1402
Qy 459 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 478
Db 1403 GCCTTCAATGGGTCTGGCACCTACTGTGTGAATTTCACTCTGGGAGATGATGCAAGCCTG 1462
Qy 479 AlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThr 498
Db 1463 GCCTCACCAGCACCTGATCTATCCCTGGCAAGACCCAGACTCCCTCTCGAGACA 1522
Qy 499 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrIleMetValThrIle 518
Db 1523 GTGAATGGGTCTGCTGATCTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582
Qy 519 LeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValVal 538
Db 1583 TTGCTGTACAAAAACACAAAGCGGTACAGACCAATAGGAACTGCCCGAGAACACCGTC 1642
Qy 539 LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAsp 558
Db 1643 AAGGGCAAGAGCGGTGAGTGTCTCTCAGCCACCGCAAGAGCCCGCTTCTCCGAGGAGAC 1702
Qy 559 ArgGluLysAspProLeuLeuGlnAspLysPro 569
Db 1703 CAGGAGAGGATCCATTGCTCCAGGACAGACCA 1735
RESULT 4
AF322054

```

```

LOCUS AF322054 2279 bp mRNA linear ROD 01-MAR-2001
DEFINITION Mus musculus dendritic cell-associated transmembrane protein
precursor (Dchil) mRNA, complete cds.
ACCESSION AF322054
VERSION AF322054.1 GI:13172896
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2279)
AUTHORS Shikano,S., Bonkobara,M., Zukas,P.K. and Ariizumi,K.
TITLE Molecular Cloning of a Dendritic Cell-Associated Transmembrane
Protein, DC-HIL, that Promotes RGD-Dependent Adhesion of
Endothelial Cells Through Recognition of Heparan Sulfate
Proteoglycans
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2279)
AUTHORS Shikano,S. and Ariizumi,K.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2000) Dermatology, University of Texas
Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
75390-9069, USA
FEATURES
source
1..2279
Location/Qualifiers
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/cell_line="XS52 DC"
/cell_type="dendritic cell"
1..2279
/gene="Dchil"
44..1768
/function="Dchil"
/function="heparan sulfate proteoglycan-dependent integrin
ligand"
/note="binds heparin; contains RGD motif; glycosylated;
ITAM: similar to Mus musculus numb product encoded by
GenBank Accession Number AJ251685"
/codon_start=1
/product="dendritic cell-associated transmembrane protein
precursor"
/protein_id="AAK14240.1"
/db_xref="GI:13172897"
/translaton="MESLGVGLLELLAAGLPLOAKRFDVYLGHQYPNHMRHNL
RQWSDENEDHLYPVWRGDKWQSGVQAVLTSQSPALVGSNITFVYNLVF
PKQEDANGNIVYEKNRDLGLTSLDHTVNTAGDGDGMDGTGRSHLRFPDRR
PPRPHGKKWSFYVFTLQYFKLGRCSARVINTVLTAGPQVMEVTVFRYGR
AYIPISKVDVVIITDQIPVFVTKSKNDRNLSDIPLRDLPIVDFLIHDPSPFLND
SAISKWNGDNTGLFVGNHNTLNHTYVNLNLTQVTAVPGPCPPSPSPPPP
SPSPSPPLPTLSTPSPSLMPTCYKSMELSDISNENCRINRYGYPRATITVEGIL
EVSIMQIADVPMPPOPANSLMDFTVCKGATPHEACTIISDPQCIAQNRVCSFVAV
DGLCLLSVRRAPNGSGTCVNTFLGDDASLALTSLTISIPKQDPDPSPLRAVGVLSI
GCLAVLVTMTVTLILYKKHKAYKPGNCPNPTVKGKLSVLISHAKAPFFRQDQKPL
IQDKPRTL"
BASE COUNT 577 a 605 c 560 g 537 t
ORIGIN
Alignment Scores:
Pred. No.: 6,44e-175 Length: 2279
Score: 2700.00 Matches: 507
Percent Similarity: 92.29% Conservative: 20
Best Local Similarity: 88.79% Mismatches: 42
Query Match: 88.21% Indels: 2
DB: 10 Gaps: 1
US-09-943-075A-2 (1-572) x AF322054 (1-2279)
Qy 1 MetClyserLeuCysGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln 20
Db 44 ATGGAAGTCTCTCGGGGCTCTGGGATTTCTGCTGCTGGCTGGAGGACTGCCCTCTCCAG 103
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40

```

Db 104 GCTGCCAAGCGATTTCGTGATGCTGGGCCATGAACGATATCCCAATCACATGAGAG 163  
 Qy 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro 60  
 Db 164 CACAACCAATTACGTGGCTGGTCTGGGATGAATAATGAATGGATGAACACCTGTATCCA 223  
 Qy 61 ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla 80  
 Db 224 GTGTGGAGGAGGAGACCGCAGGTGGAGGACTCCTGGGAGGAGGCGGTGTGCAGGCA 283  
 Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
 Db 284 GTCTGACCACTGACTCACCGGCTGTGGTGTCCATATATCACCTTGTGTGTGAACCTG 343  
 Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
 Db 344 GTGTTCCTCCAGATGCCAGAGGAAGATGCTAATGGCAATATCGTCTATGAGAAGAACTGC 403  
 Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140  
 Db 404 AGGAATGATTGGGACTGACCTCTGACCTGCATGCTACAACTGCAGCTGCAGGGGCAGAT 463  
 Qy 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
 Db 464 GATGGTCACTGGGAAGATGGCACCGGAGGACGATCTCAGGTTCGGGACAGGAGG 523  
 Qy 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
 Db 524 CCCTTCCTCGCCCCCATGATGGAGAGAAATGGAGCTTGTCTACGCTTTCACACACTT 583  
 Qy 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
 Db 584 GGCCAGTATTCCAAAACCTGGGTGCTGTTACGACGCGGTCTTATAAACACAGCTCAAC 643  
 Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220  
 Db 644 TTGACAGCTGGCCCTCAGGTCATGGAGGTGACTGTCTTTCGAAGATACCGCGGCGCATAC 703  
 Qy 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
 Db 704 ATTCCTCATCTCGAAGGTGAAGATGTGTATGTATGAACAGATCAGATCCCTGTATTCGTG 763  
 Qy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
 Db 764 ACCATGTCCCAAGAAGATGACAGGAACCTTGTCTGATGAGATCTTCTCAGACACCTCCCC 823  
 Qy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
 Db 824 ATCGTCTTCGATGTCTCATTCATGATGCCAGCCACTTCTCTCAACGACTCTGCCATTCC 883  
 Qy 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
 Db 884 TACAAGTGAACCTTGGGGACACACTGGCTGTGTGCTCCCAACATCACACTTGAAT 943  
 Qy 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
 Db 944 CACACTTATGTGCTCAATGGAACCTTCAACCTTAACTTCACTCAGCGTGAACCTGCGCC 1003  
 Qy 321 GlyProCysProSerProThrProSerProSerSerThrSerProSerPro 338  
 Db 1004 GGGCATGCGCTCCCTCCCTTCGCTTCGACTCCGCTCCACCTTCACTCCGCTCACCT 1063  
 Qy 339 AlaSerSerProSerProThrLeuSerThrProSerProSerSerLeuMetProThrGlyTyr 358  
 Db 1064 CCGCCCTCACCTCTGCCACCATATCAACACTAGCCCTCTTAAATGCTACTGCTTAC 1123  
 Qy 359 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 378  
 Db 1124 AAATCCATGGAGCTGACGTGACATTTCAATGAAAACCTGCGGAATAAACAGATATGCGCTAC 1183  
 Qy 379 PheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAla 398

Db 1184 TTCAGAGCCACCACCAATTTGTAGAGGGGATCTCGGAAGTCAAGTCAATCATCATGACATAGCA 1243  
 Qy 399 AspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCys 418  
 Db 1244 GATGTCCCATGCCCCACACCGCAGCTGCCAACTCCCTGATGGACTTCACTGTGACCTGC 1303  
 Qy 419 LysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 438  
 Db 1304 AAAGGGCCACCCCATGGAAGCCTGTAGATCATCTCCGACCCCACTGCCAGATGCC 1363  
 Qy 439 GlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArg 458  
 Db 1364 CAGAACCGGGTCTGCAGCCCTGTGGCTGTGGATGGCTGTGCCTGCTCTGTGTGAGAAGA 1423  
 Qy 459 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeu 478  
 Db 1424 GCCTTCAATGGGTCTGGCACCTACTGTGTGAATTTCACTCTGGGAGATGATGCAAGCCTG 1483  
 Qy 479 AlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThr 498  
 Db 1484 GCCTTCACAGCACCCCTGATCTATCCCTGGCAAGACCCAGACTCCCTCTGTGAGAGCA 1543  
 Qy 499 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIle 518  
 Db 1544 GTCAATGGTGTCTGATCTCCATGGCTGCTGCTGTGTTCACCATGTTTACCATC 1603  
 Qy 519 LeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValVal 538  
 Db 1604 TTGCTGTACAAAAACACCAAGCGTACAAAGCAATAGGAAACTGCCCGCAACACGCTC 1663  
 Qy 539 LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAsp 558  
 Db 1664 AAGGGCAAGCGCTGAGTGTCTCTCAGCCAGCGGAAGCCCGTCTTCTCGAGAGAC 1723  
 Qy 559 ArgGlyLysAspProLeuGlnAspLysPro 569  
 Db 1724 CAGGAGAAGGATCCATTGCTCCAGGACAAGCCA 1756  
 RESULT 5  
 BC026375 2299 bp mRNA linear ROD 07-AUG-2002  
 LOCUS Mus musculus, glycoprotein (transmembrane) nmb, clone MGC:31121  
 DEFINITION IMAGE:4164706, mRNA, complete cds.  
 ACCESSION BC026375  
 VERSION BC026375.1 GT:20071338  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 2299)  
 Strausberg, R.  
 Direct Submission  
 Submitted (02-APR-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcdpaxil.stanford.edu](mailto:mcdpaxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.  
 REMARK  
 COMMENT  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 43 Row: d Column: 20







```
Db 1130 CTGGAGCTGAGTAGGATTCCTGTGTAAGAACTGCCAGATTAACAGATATGCGCCACTTTCAA 1189
QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400
Db 1190 GCCACCATCACAATTTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTC 1249
QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValIleThrCysLysGly 420
Db 1250 CTGATGCCGGTCCGATGCCCTGGAAGCTCCCTAATAGACTTTCTCGTGACCTGCCCAAGGG 1309
QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440
Db 1310 AGCATTCCTCCAGGAGCTGTGTACATCATTTCTGACCCACCTGGAGATCACCCAGAAC 1369
QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460
Db 1370 ACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGCTGACTGTGAGACGAACCTTC 1429
QY 461 AsnGlySerGlyThrTyrcysValAsnPheThrLeuGlyAspAlaSerLeuAlaLeu 480
Db 1430 AATGGTCTGGACGTACTGTGAACCTCACCTGGGGATGACACAAAGCTGGCTCTC 1489
QY 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500
Db 1490 ACGAGCACCCGTATTTCTGTCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAC 1549
QY 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520
Db 1550 AGTCCCTGTATCTCGTGGCTGCTGGCCATATTTGTCACGTGATCTCCCTGTGGTG 1609
QY 521 TyrLysLysHisLysThrTyrcysProIleGlyAsnCysThrArgAsnValValLysGly 540
Db 1610 TACAAAAACACAGGAATACACCCCAATAGAAATAGTCTGGGAATGTGTCAGAAAGC 1669
QY 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560
Db 1670 AAAGCCTGAGTGTCTTCTCACCCTGCAAGGCGCTGTTCTCCGGGAACACGAA 1729
QY 561 LysappProLeuLeuGlnAspLys 568
Db 1730 AAGATCCGCTACTCAAAACCAA 1753

RESULT 8
LOCUS AR018814 2669 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 101 from patent US 5783182.
ACCESSION AR018814
VERSION AR018814.1 GI:3973928
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2669)
AUTHORS Thompson,T.C.
TITLE Method for identifying metastatic sequences
JOURNAL Patent: US 5783182-A 101 21-JUL-1998;
FEATURES
source Location/Qualifiers
1..2669
/organism="unknown"
BASE COUNT 752 a 595 c 597 g 725 t
ORIGIN

Alignment Scores:
Pred. No.: 3,14e-138 Length: 2669
Score: 2163.00 Matches: 394
Percent Similarity: 82.57% Conservative: 75
Best Local Similarity: 69.37% Mismatches: 85
Query Match: 70.66% Indels: 14
DB: 6 Gaps: 1

US-09-943-075A-2 (1-572) x AR018814 (1-2669)
QY 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGln 20
```

```
Db 92 ATGGAAATGCTCTACTATTCTCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGAT 151
QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrcysProAspHisMetArgGlu 40
Db 152 GCCGCCAAACGATTTCATGATGCTGGCAATCAAGAGACCTTCTGCTTACATGAGGGAG 211
QY 41 AsnAsnGlnLeuArgGlyTyrcysSerAspGluAsnGlnTyrcysProAspGluGlnLeuTyrcys 60
Db 212 CACAATCAATTAATGGCTGCTTCTGATGAAATGACTGGAATGAAATGCTTACCA 271
QY 61 ValTyrArgGlyGlyGlyLysArgTrpLysAspSerTrpGluGlyArgValGlnAla 80
Db 272 GTGTGAAGCGGGAGACATGAGGTGGAAAAAACCCTGGAGAGGAGCCCTGTGAGGGG 331
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 332 GTCTTGACCAAGTCACTACCAGCCCTGCTGGGCTCAATATATAACATTGGCGGTGAACCTG 391
QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrcysGluArgAsnCys 120
Db 392 ATATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAAGAACTGC 451
QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrcysValTyrcysAsnTrpThrGlyAlaAsp 140
Db 452 AGAAATGAGGCTGGTATTCTGCTGATCCATATCTTTACAACCTGGACAGCATGGTCAGAG 511
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
Db 512 GACAGTACGGGGAATGGCAGCCGCAAGCCATCAACGCTTCTCCCTGTATGGGAAA 571
QY 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrcysValPheHisThrLeu 180
Db 572 CCTTTCTCACCACCCCGGATGGAGAGATGGAATTTCTATCTACGCTTCTCCACACACTT 631
QY 181 GlyGlnTyrcysPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Db 632 GGTCAAGTATTTCCAGAAATGGGACGATGTTCTGAGAGAGTTTCTGTGAACACAGCAAT 691
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyrcys 220
Db 692 GTGACACTTGGCGCTCAACTCATCGAAGTGAAGTCTGTACAGAAAGACATGCGCGGCATAT 751
QY 221 IleProIleSerLysValLysAspValTyrcysValIleThrAspGlnIleProIlePheVal 240
Db 752 GTTCCCATCGCAGTAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTGTG 811
QY 241 ThrMetTyrcysGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
Db 812 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGCAAAACCTTCTCCAAAGATCTCCCC 871
QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrcysSerAlaIleSer 280
Db 872 ATTTATGTTTGTATGCTCTGATTCATGATCCTAGCCACTTCTCTCAATATTATCTACCAAT 931
QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
Db 932 TACAAGTGGAGTTCGGGGATAATACTGGCTGTTTGTTCACCAATCATATCTGTAAT 991
QY 301 HisThrTyrcysValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
Db 992 CACAGTATGTCTCAATGGAAACCTTCAGCCTTAAACCTCACTGTGAAGCTGCAGACCA 1051
QY 321 GlyProCysProSerProThrProSerProSerSerSerSerSerSerProAlaSer 340
Db 1052 GGACCTTGTCCGCCACCCACCCACCCACCCACCCACCT----- 1087
QY 341 SerProSerProThrLeuSerThrProSerProSerProSerProSerProSerProSer 360
Db 1088 -----TCAAAACCCACCCCTTCTTTAGGACCTGCTGTGACAAACCC 1129
QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrcysTyrcysPheArg 380
Db 1129 -----TCAAAACCCACCCCTTCTTTAGGACCTGCTGTGACAAACCC 1129
```



```

Db 1130 CTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTACAGATATGGCCACTTTCAA 1189
Qy 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGluValAlaAspVal 400
Db 1190 GCCACCATCAAACTGTAGAGGAATCTTAGAGTTTAACATCATCCAGATCAGACAGCTC 1249
Qy 401 ProIleProThrLeuGlnProAsnSerLeuMetAspPheIleValThrCysLysGly 420
Db 1250 CTGATGCCGGTCCATGCCCTGAAAGCTCCCTAATAGACTTTGTCTGCTGACCTGCCAAGG 1309
Qy 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440
Db 1310 AGCATTCCACGGAGTCTGTACATCATTTCTGACCCACCTCGGAGATCACCAGAAC 1369
Qy 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460
Db 1370 ACAGTCTCGACCCCTGTGGATGTGGATGAGATGTCTGCTGCTGCTGAGACGAACCTTC 1429
Qy 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerIleValAlaLeu 480
Db 1430 AATGGTCTGGAGCTACTGTGTGAACCTCACCTGGGGATGACACAGCCTGGCTCTC 1489
Qy 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500
Db 1490 ACGAGCACCTGATTTCTGTTCTCGACAGAGACCCAGCCTCGCTTTAAGGATGGCAAC 1549
Qy 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520
Db 1550 AGTGCCTCGATCTCGTTGGCTGCTTGGCCATATTTGTACATGTGATCTCCCTGTGTG 1609
Qy 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540
Db 1610 TACAAAAACAACAAGGAATACACCCCAATAGAAAATAGCTCTGGGAATGTGTGTCAGAAC 1669
Qy 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560
Db 1670 AAAGCCCTGAGTGTCTTCTCAACCGTGCANAAAGCCGTGTCTCTCCGGGAAACAGGAA 1729

Qy 561 LysAspProLeuLeuGlnAspLys 568
Db 1730 AAGGATCGGCTACTCAAAAACCAA 1753

RESULT 9
HSNMB HSNMB 2669 bp mRNA linear PRI 09-FEB-1995
LOCUS H.sapiens NMB mRNA.
DEFINITION H.sapiens NMB mRNA.
ACCESSION X76534
VERSION X76534.1 GI:666042
KEYWORDS NMB gene.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2669)
AUTHORS Weterman,M.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1993) M. Weterman, University of Nijmegen, Dept
of Biochemistry, PO Box 9101, 6500 HB Nijmegen, NETHERLANDS
REFERENCE 2 (bases 1 to 2669)
AUTHORS Weterman,M.A., Ajubi,N., van Dinter,I.M., Degen,W.G., van
Muljen,G.N., Ruitter,D.J. and Bloemers,H.P.
TITLE nmb, a novel gene, is expressed in low-metastatic human melanoma
cell lines and xenografts
JOURNAL Int. J. Cancer 60 (1), 73-81 (1995)
MEDLINE 95113576
PUBMED 7814155
FEATURES
Location/Qualifiers
1..2669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NMB"
/cell_line="MV1"
/clone_lib="MV1"

```

```

gene 92..1774
CDS /gene="NMB"
92..1774
/gene="NMB"
/codon_start=1
/protein_id="CAA54044.1"
/db_xref="GI:666043"
/db_xref="SWISS-PROT:Q14956"
/translacion="MECLYVFLGLLAAARLPDAAKRFDHVLGNRPSSAYMRHNL
NGSSDENNEKLYPMKRGDMKNSKGRVOAVLTSDSPALVGSNITFAVNLIF
PRCKEDANGNIYKNCNEACLADPYVYNTKNSDSENGENTGSHNVFDPGK
FPHPHQRWNFIYFHLQIFQKLRCSVRVSNANTANFLGQLMEVTVYRRHGK
STINYKNSFGDNTGLFVSTNHNTHYVLNGTFLSNLTKAAAPGPCPPPPPPSK
PTSLGDPAGNPLESRIPDENCIINRYHQATITIVEGILEVINIQMTDVLMPVM
PESSLIIDFVVTGCGSIPTVECTIISDPTCEITQNTVCSVDENCLITVRRTPNGS
TYCWNLTGDDTSLAULTSLISVDRDPASPLRMANSALISVGLAIFVTVISLVLVYK
KHKEYNPENSPGNVVRSGLSVFLNRAKAVFPFGNQEKDPLLKQEFKYS"
mat_peptide 155..1771
/gene="NMB"
BASE COUNT 752 a 595 c 597 g 725 t
ORIGIN
Alignment Scores:
Pred. No.: 3,14e-138 Length: 2669
Score: 2163.00 Matches: 394
Percent Similarity: 82.57% Conservative: 75
Best Local Similarity: 69.37% Mismatches: 85
Query Match: 70.66% Indels: 14
DB: 9 Gaps: 1
US-09-943-075A-2 (1-572) x HSNMB (1-2669)
Qy 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20
Db 92 ATGGAATGTCCTACTATTTCTCTGGGATTTCTGCTCTGCTGCAAGATTGTCACCTGTAT 151
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
Db 152 GCCGCCAAACGATTTCATGATGCTGGCAATCAAGAACCTTCTGCTTACATGAGGGAG 211
Qy 41 AsnAsnGlnLeuArgGlyTyrSerSerAspGluAsnGlnTyrPaspGluGlnLeuTyrPro 60
Db 212 CACAATCAATTAATGCGTGTCTCTCATGATGAAATGACTGGAATGAAACCTCTACCCA 271
Qy 61 ValTyrArgArgGlyGluGlyArgTyrLysAspSerTyrGluGlyArgValGlnAla 80
Db 272 GTGTGGAAGCGGGAGACATCAGGTGGAAAAAATCTCTGGAAGGAGCGCGTGCAGCG 331
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 332 GTCCGTGACAGTGACTCACCAGCCCTCGTGGGCTCAAAATATAACATTTGCGGTGAACCTG 391
Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120
Db 392 ATATTCCTCTAGATGCCAAAGAAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 451
Qy 121 ArgSerAspLeuLeuAlaSerAspProTyrValTyrAsnTyrThrThrGlyAlaAsp 140
Db 452 AGAAATGAGCGTGGTATTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG 511
Qy 141 AspGluAspTyrPheGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
Db 512 GACATGACGGGGAATAATGGCACCCGCCAAAGCCATCAACGCTCTTCCTGATGGGAAA 571
Qy 161 ProPheProArgProHisGlyArgLysLysTyrPasnPheValTyrValPheHisThrLeu 180
Db 572 CCTTTTCTCACCACCCCGGATGGAGAGATGGAAATTTTATCTACGCTTCTCCACACACTT 631
Qy 181 GlyGlnTyrPheGlnLysLeuGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Db 632 GGTCAGTATTTCCAGAAATTTGGACGATGTTTCAGTGAGAGCTTCTCTGTAACACAGCCAT 691

```

QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220  
Db 692 GTGACACTTGGCGCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 751  
QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
Db 752 GTTCCCATCGCAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCCTGTGTGTGG 811  
QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
Db 812 ACTATGTTCCAGACAGATCGAAATTCATCCGACGAAACCTCTCTCAAGATCTCCCC 871  
QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
Db 872 ATTATGTTGATGCTCTGATTCATGATCTCTAGCCACTTCTCAATATTATTTACCATAT 931  
QY 281 TyrLysTyrAspPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
Db 932 TACAAGTGGAGCTTCGGGGAATACTGGCTGTGTTTCCACCAATCATATCTGTGAT 991  
QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
Db 992 CACACATATGCTCAATGGAACTCTAGCCCTTAACCTCACTGTGAAGCTGACGACCA 1051  
QY 321 GlyProCysProSerProThrProSerProSerSerSerThrThrSerProSerProAlaSer 340  
Db 1052 GGACCTGTGCGCCACGCCACCACCACCCAGACCT----- 1087  
QY 341 SerProSerProThrLeuSerThrProSerProSerProSerLeuMetProThrGlyTyrLysSer 360  
Db 1088 -----TCAAAACCCACCCCTTCTTAGGACCTGCTGGTGACAACCC 1129  
QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380  
Db 1130 CTGGAGCTGAGTAGATTCCTCTGTAAGAACTGCCAGATTAACAGATATGCCACACTTCAA 1189  
QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400  
Db 1190 GCCACCATCACAAATGTAGAGGAATCTTAGAGGTAAACATCATCCAGATGACAGACGTC 1249  
QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysIysGly 420  
Db 1250 CTGATCGCGTCCCATGGCTGAAAGCTCCCTAATAGACTTGTGCTGACCTGCCAAGG 1309  
QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440  
Db 1310 AGCATTCACCGAGGCTGTACCATCATTTCTGACCCCACTCGAGATCAACCCAGAAC 1369  
QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArgAlaPhe 460  
Db 1370 ACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGTGACTGTGAGACGAACTTC 1429  
QY 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeuAlaLeu 480  
Db 1430 AATGGGTCTGGACACTACTGTGAACTCAACCTGGGGGATGACAAAGCTGGCTCTC 1489  
QY 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500  
Db 1490 ACGAGCAGCCCTGATTCCTGCTCAGACAGACCCACCTCGCTTTAGGATGGCAAC 1549  
QY 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
Db 1550 AGTGGCCCTGATCTCGTGTGGCTGTGGCCATATTGTGACTGTGATCTCCCTCTCTGGTG 1609  
QY 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540  
Db 1610 TCAAAAAACACAGGAATAACCCCAATAGAAAAATAGTCTCGGGAATGTGGTCAAGAC 1669  
QY 541 LysGlyLeuSerValPheIleLysSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560  
Db 1670 AAAGGCTTGAGTGCTTCTCAACCGTCAAAAGCCGTGTTCTTCCCGGGAACACAGGAA 1729

QY 561 LysAspProLeuLeuGlnAspLys 568  
Db 1730 AAGCATCGCTACTCAAAAACCAA 1753  
RESULT 10  
AX358788 2683 bp DNA linear PAT 13-FEB-2002  
LOCUS Sequence 41 from Patent WO0193983.  
DEFINITION AX358788  
ACCESSION AX358788  
VERSION AX358788.1 GI:18675295  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Baker,K.P., Desnoyers,L., Gerritsen,M.E., Goddard,A.,  
Godowski,P.J., Grimaldi,J.C., Gurney,A.L., Smith,V., Stephan,J.P.,  
Watanabe,C.K. and Wood,W.I.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding  
the same  
JOURNAL Patent: WO 0193983-A 41 13-DEC-2001;  
Genentech Inc. (US)  
FEATURES  
source 1..2683  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 751 a 604 c 596 g 732 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,7e-138 Length: 2683  
Score: 2162.00 Matches: 396  
Percent Similarity: 82.72% Conservative: 78  
Best Local Similarity: 69.11% Mismatches: 87  
Query Match: 70.63% Indels: 12  
Gaps: 2  
US-09-943-075a-2 (1-572) \* AX358788 (1-2683)  
QY 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20  
Db 82 ATGGAATGTCTCTACTATTTCTGGGATTTCTGCTCCCTGCTGCAAGATTGCCACTGTAT 141  
QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
Db 142 GCCGCCAAACGATTTCTATGATGTCTGGGCAATGAAGACCTTCTGCTTACATGAGGAG 201  
QY 41 AsnAsnGlnLeuArgGlyTyrPheSerSerAspGluAsnGluTyrAspGluGlnLeuTyrPro 60  
Db 202 CACAATCAATTAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACCTCTACCCA 261  
QY 61 ValTyrArgArgGlyClyGlyArgTyrPheSerSerTrpGluGlyGlyValGlnAla 80  
Db 262 GTCTGAAAGCGGGAGACATAGAGTGGAAACACTCCTGGAAGGAGGCCGTGTGACGGCG 321  
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
Db 322 GTCTTGACCACTGACTACACAGCCCTCGTGGGCTCAAAATATAACATTTGCGGTGAACCTG 381  
QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
Db 382 ATATTCTCTAGATGCCAAAGGAAGATGCCAATGCAACATAGTCTATGAGAAGAACTGC 441  
QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrThrGlyAlaAsp 140  
Db 442 AGAATAGCGGTGTTTATCTCTGCTGATCGGTATGTTTACAACTGGACAGCATGCTCAGAG 501  
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
Db 502 GACAGTGACGGGGGAAAAATGGCCACCGCCAAAGCCCATCAACGCTCTTCCCTGTATGGGAAA 561  
QY 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180









```

Db 942 GGAAGTGGCTTATTGTAGAAAGCGCGCTACTACGCTCTCACACCTCTCTCTGCAAGA 1001
Qy 308 ThrPheAsnPheAsnLeuThrValGlnThrAlaValProGlyProCysProSerProThr 327
Db 1002 AACTTCACCCCTGAATTTTAACTGTTCAAGCTATTATACCTGTGCTTGCAGGCGAGTAACA 1061
Qy 328 ProSerProSerSerThrSerProSerProAlaSerSerProAlaSerSerProThrLeuSer 347
Db 1062 CCCACTCCA-----TCACCTGCCACCCCGACGAGTAACG 1094
Qy 348 ThrProSerProSerLeuMetProThrGlyTyrLysSerMetGluLeuSerAspIleSer 367
Db 1095 ACTGATGCATCTTCTAATTTCAGACCCCTCTGCTCCCAATGAGATGGCAGAGAACAATCCT 1154
Qy 368 AsnGluAsnCysAlaGleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAsp 387
Db 1155 GATGGAGCGTCCCATNTTACAGATATGGATACACTACACGCTGGTATACCACTTGTAGAG 1214
Qy 388 GlyIleLeuGluValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnPro 407
Db 1215 GGAATCCTTGAGTGAATAATTATCCAGATGACAAGCATCCAGATGACAGAAAGCCAGCT 1274
Qy 408 AspAsnSerLeuMetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCys 427
Db 1275 GAAATCCACTGGTGGACTTGTGTGTACCTGCCAAGGAGTTTCCTACAGATGTCTGT 1334
Qy 428 ThrIleIleSerAspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAla 447
Db 1335 ACACGAGTTTTCAGCCACGCTGCCAGGTGTCTCAGGGCATGGTATGATCCCGTGGTC 1394
Qy 448 ValAspGluLeuCysLeuLeuSerValArgAlaPheAsnGlySerGlyThrTyrCys 467
Db 1395 GTCACTGTAATGCGTACTACACATGAAGGAGAGCTTTTGATGAACCTGGAACGTACTGT 1454
Qy 468 ValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIle 487
Db 1455 ATAATATACCTTGGGGGATGACACCGCAAGCCCTTGCACGCGCACTTATTCTGTGA 1514
Qy 488 ProGlyLysAspLeuGlySerProLeuArgThrValAsnGlyValLeuIleSerIleGly 507
Db 1515 AATGGA-----GGATCATCTCAGGGGACAAAGGCTCTCTTCATCTTTCTGGA 1565
Qy 508 CysLeuAlaMetPheValThrMetValThrIleLeuLeuTyrLysLysHisLysThrTyr 527
Db 1566 TTGCTGGCAGTGTGTGCTGCTTGGGCGCTTTTGTCTTTTACAAGAGATACAAACAATAC 1625
Qy 528 LysProIleGlyAsnCysThrArgAsnValValLysGlyLysGlyLeuSerValPheLeu 547
Db 1626 AAACCTATTGAGAGAAGTGGGGACAGCAGACAGACAGGAGGAGGAGTGTGCTATGTGC 1685
Qy 548 SerHisAlaLysAlaProPheSerArgGlyAspArgGluLysAspProLeuLeuGlnAsp 567
Db 1686 AGCAACTTCAAGAGCCCTTTTCTCCCTAAGACACTGAGAGAAATCTCTGTGTGAAGAGC 1745
Qy 568 LysPro 569
Db 1746 AAACCA 1751

```

```

RESULT 14
LOCUS AR018825 620 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 113 from patent US 5783182.
ACCESSION AR018825
VERSION AR018825.1 GI:3973939
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 620)
AUTHORS Thompson,T.C.
TITLE Method for identifying metastatic sequences
JOURNAL Patent: US 5783182-A 113 21-JUL-1998;

```

```

FEATURES             Location/Qualifiers
     source            1..620
BASE COUNT           143 a 207 c 128 g 142 t
ORIGIN
Alignment Scores:
Pred. No.:           9.57e-56      Length: 620
Score:               941.00        Matches: 183
Percent Similarity: 91.75%        Conservative: 6
Best Local Similarity: 88.83%      Mismatches: 15
Query Match:         30.74%        Indels: 3
DB:                  6            Gaps: 1

US-09-943-075A-2 (1-572) x AR018825 (1-620)
Qy 274 LeuAsnTyrSerAlaIleSerTyrLysTyrAsnPheGlyAspAsnThrGlyLeuPheVal 293
Db 2 CTCGAACGACTCTGCCATTTCTACAAGTGAACCTTTGGGACAACTGGCCTGTTTGTGTC 61
Qy 294 SerAsnAsnHisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeu 313
Db 62 TCCAAACATCACACTTTGAATCACACTTATGCTCAATGGAACCTTCAACCTTAACCTC 121
Qy 314 ThrValGlnThrAlaValProGlyProCysProSerProThrProSer-----ProSer 331
Db 122 ACCGTGCAAACTGCAGTGGCCGGCCATGCCCTCCCTTCGCTTCGACTCGGCCTCCA 181
Qy 332 SerSerThrSerProSerProAlaSerSerProSerProThrLeuSerThrProSerPro 351
Db 182 CCTTCAACTCGGCCCTCACCTCGCCCTCACCTCTGCCACATTATCAACACCTAGCCGC 241
Qy 352 SerLeuMetProThrGlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnCys 371
Db 242 TCTTTAATGCTTACTGTTTACAAATCCATGAGCTGAGTGACATTTTCCAATGAAACATGC 301
Qy 372 ArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGlu 391
Db 302 CGAATAAACAGATATGCTACTTTCAGAGCCACCATCACAAATGTAGAGGGGATCTGGNA 361
Qy 392 ValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeu 411
Db 362 GTCAGCATCATGTAGATAGCAGATGTCCCATGATGCCACACCCGAGCTGCCAATCCCTG 421
Qy 412 MetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSer 431
Db 422 ATGCACTTCACTGTGACCTGCAAGGGGCGCCCATGGAAGCCCTGTACGATCATCTCC 481
Qy 432 AspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeu 451
Db 482 GACCCACCTGCCAGATCGCCCAAGACCGGGTCTGCAGCCCTGTGGCTGTGATGGGCTG 541
Qy 452 CysIleLeuSerValArgArgAlaPheAsnGlySerGlyThrTyrCysValAsnPheThr 471
Db 542 TGCCTG-CTGTCTGTGAGAAGAGCCTTCAATGGGTCTGGCACCTACTGTGTGAATTTCACT 600
Qy 472 LeuGlyAspAspAlaSer 477
Db 601 CTGGGAGATGATGCAAGC 618

```

```

RESULT 15
LOCUS AC098114 162541 bp DNA linear HTG 12-JUL-2002
DEFINITION Rattus norvegicus clone CH230-38F9, *** SEQUENCING IN PROGRESS ***,
ACCESSION AC098114 72 unordered pieces.
VERSION AC098114.4 GI:21729266
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

REFERENCE  
AUTHORS

1 (bases 1 to 162541)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbraia,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowie,S., Brieva,M., Brown,M., Bryant,N.P.,  
 Buay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,W., Havlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,  
 Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,  
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokwenkwo,S., Oguh,M., Okwuonwu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,  
 Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 162541)  
 Worley,K.C.  
 Direct Submission  
 Submitted (23-OCT-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 162541)  
 Worley,K.C.  
 Direct Submission  
 Submitted (12-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 11, 2002 this sequence version replaced gi:17969936.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GGIC  
 Center clone name: CH230-38P9  
 ----- Summary Statistics  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 106765 bases at least Q40  
 Consensus quality: 110840 bases at least Q30  
 Consensus quality: 113567 bases at least Q20  
 -----

\* NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 72 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1201: contig of 1201 bp in length  
 1202 1301: gap of unknown length  
 1302 2388: contig of 1087 bp in length  
 2389 2488: gap of unknown length  
 2489 3513: contig of 1025 bp in length  
 3514 3613: gap of unknown length  
 3614 4731: contig of 1118 bp in length  
 4732 4831: gap of unknown length  
 4832 5962: contig of 1031 bp in length  
 5963 7083: contig of 1121 bp in length  
 7084 7183: gap of unknown length  
 7184 8559: gap of unknown length  
 8560 10098: contig of 1539 bp in length  
 10099 10198: gap of unknown length  
 10199 11258: contig of 1060 bp in length  
 11259 11358: gap of unknown length  
 11359 12491: contig of 1133 bp in length  
 12492 12591: gap of unknown length  
 12592 13632: contig of 1041 bp in length  
 13633 13732: gap of unknown length  
 13733 14876: contig of 1144 bp in length  
 14877 16015: contig of 1039 bp in length  
 16016 16115: gap of unknown length  
 16116 17334: contig of 1219 bp in length  
 17335 17435: gap of unknown length  
 17436 18762: contig of 1328 bp in length  
 18763 18862: gap of unknown length  
 18863 20046: contig of 1184 bp in length  
 20047 20446: gap of unknown length  
 20447 21488: contig of 1342 bp in length  
 21489 22158: contig of 1127 bp in length  
 22159 22815: gap of unknown length  
 22816 23925: contig of 1110 bp in length  
 23926 24025: gap of unknown length  
 24026 25336: contig of 1511 bp in length  
 25337 25637: gap of unknown length  
 25638 26722: contig of 1086 bp in length  
 26723 26822: gap of unknown length  
 26823 27978: contig of 1156 bp in length  
 27979 28078: gap of unknown length  
 28079 29407: contig of 1329 bp in length  
 29408 29507: gap of unknown length  
 29508 30541: contig of 1034 bp in length  
 30542 31898: contig of 1257 bp in length  
 31899 31998: gap of unknown length  
 31999 33108: contig of 1110 bp in length  
 33109 33208: gap of unknown length  
 33209 34411: contig of 1203 bp in length  
 34412 34511: gap of unknown length  
 34512 36645: contig of 2134 bp in length  
 36646 36745: gap of unknown length  
 36746 38188: contig of 1443 bp in length  
 38189 38288: gap of unknown length  
 38289 39479: contig of 1191 bp in length  
 39480 39579: gap of unknown length  
 39580 41328: contig of 1749 bp in length  
 41329 41428: gap of unknown length  
 41429 42540: contig of 1112 bp in length  
 42541 42640: gap of unknown length







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 10:06:22 ; Search time 356 Seconds  
(without alignments)  
3618.379 Million cell updates/sec

Title: US-09-943-075A-2  
Perfect score: 3061  
Sequence: 1 MESLQGVFLVLLAAGLPQ.....PFSRGDRKDPDLLQKPMWL 572

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Maximum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framet.p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09943075/runat\_07042003\_155630\_17920/app\_query.fasta\_1.711  
-DB=N\_Geneseq\_101002 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -DOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09943075.ecgn\_1\_1\_263\_@runat\_07042003\_155630\_17920 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3040	99.3	2303	19	AAT96034 Rat kidney injury
2	2700	88.2	2213	21	AAA96729 Polynucleotide iso
3	2163	70.7	2669	18	AAT69318 Murine metastatic
4	2163	70.7	2669	18	AAT69328 Murine metastatic
5	2163	70.7	2669	21	AAT55715 Human NRB cDNA seq
6	2163	70.7	2669	24	ABO88185 Human osteoblast d
7	2162	70.6	2683	24	ABK33556 cDNA encoding huma
8	941	30.7	620	18	AAT69340 Murine metastatic
9	729	23.8	524	18	AAT69342 Murine metastatic
10	717	23.4	473	18	AAT69342 Murine metastatic
11	627.5	20.5	2115	16	AAO96055 Sequence encoding
12	627.5	20.5	2130	22	AAH43500 Human melanoma ant
13	627.5	20.5	2130	24	AA514396 cDNA encoding huma
14	627	20.5	2172	16	AAT02716 MART-1 melanoma an
15	627	20.5	2172	22	AA345525 DNA encoding Melan
16	622.5	20.3	1986	22	AAD07346 Modified tumour-as
17	622.5	20.3	1986	22	AAD07346 Modified tumour-as
18	622.5	20.3	1986	22	AAH20209 Human gpi100m nucle
19	622.5	20.3	2534	22	AAH20209 Modified gpi100m en
20	622	20.3	2131	18	AAH20209 Modified gpi100m en
21	621	20.3	2131	16	AAH20209 Modified gpi100m en
22	613.5	20.0	169739	24	ABO88186 Melanoma-specific
23	564.5	18.4	755	22	AAH23139 Human osteoblast d
24	524	17.1	335	18	AAT69350 Murine metastatic
25	524	17.1	335	18	AAT69350 Murine metastatic
26	524	17.1	335	18	AAT69321 Murine metastatic
27	524	17.1	335	18	AAT69321 Murine metastatic
28	504	16.5	354	18	AAH69341 Murine metastatic
29	471	15.4	286	18	AAT69348 Murine metastatic
30	462.5	15.1	318	22	ABA47986 Human breast cell
31	462.5	15.1	318	22	ABA65873 Human foetal liver
32	462.5	15.1	318	22	ABA32959 Probe #11425 for g
33	462.5	15.1	318	22	ABA41284 Human brain expres
34	462.5	15.1	318	22	AAK40016 Human bone marrow
35	462.5	15.1	318	22	AAK40016 Human bone marrow
36	462.5	15.1	318	22	AAI20805 Probe #10738 for g
37	462.5	15.1	318	22	AAI46035 Probe #14721 used
38	462.5	15.1	318	22	AAI06513 Probe #6504 used t
39	427	13.9	262	18	AAH69320 Human genome-deriv
40	418	13.7	309	18	AAT69333 Murine metastatic
41	409	13.4	260	18	AAT69333 Murine metastatic
42	405.5	13.2	478	22	ABA42857 Human breast cell
43	405.5	13.2	478	22	ABA53280 Human foetal liver
44	405.5	13.2	478	22	ABA23054 Probe #1520 for ge
45	405.5	13.2	478	22	AAK01537 Human brain expres

ALIGNMENTS

RESULT 1  
AAT96034  
ID AAT96034 standard; cDNA; 2303 BP.  
XX AAT96034;  
XX  
XX 21-MAY-1998 (first entry)  
XX  
XX Rat kidney injury related molecule (KIM) cDNA clone 4-7.  
DE  
KW Kidney injury related molecule; KIM; rat; renal disease; injury;  
KW nephritis; tissue regeneration; therapy; ss.  
XX  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 107..1825  
/\*tag= a

XX WO9744460-A1.  
 XX 27-NOV-1997.  
 XX 23-MAY-1997; 97WO-US09303.  
 XX 23-AUG-1996; 96US-0023442.  
 XX 24-MAY-1996; 96US-0018228.  
 XX (BIOJ ) BIOGEN INC.  
 XX Ronventre JV, Cate RL, Hession CA, Ichimura T, Sanicola-Nadel M;  
 PI Wei H;  
 XX WPI; 1998-018514/02.  
 DR P-PSDB; AAW38335.  
 XX  
 PT DNA encoding kidney injury related molecule - which is upregulated  
 PT in injured or regenerating tissue, useful to promote growth of new  
 PT tissue and survival of damaged tissue  
 XX  
 PS Claim 1; Page 38-41; 68pp; English.  
 XX  
 CC CDNA clone 4-7, deposited as ATCC 98062, codes for a rat protein  
 CC (see AAW38335), designated kidney injury related molecule (KIM), that  
 CC is up-regulated in injured or regenerating tissue. Representational  
 CC difference analysis was used to examine cDNA libraries prepared from  
 CC ischaemic and normal rat adult kidneys. 3 Cycles of hybridisation  
 CC and selective amplification gave 3 fragments present only in the  
 CC injured kidney library. Screening of the injured kidney library  
 CC with these fragments gave clone 4-7, as well as clones 3-2 (see  
 CC AAT96032) and 1-7 (see AAT96033), which are splice variants encoding  
 CC an identical KIM polypeptide (see AAW38334). A labelled probe from  
 CC clone 1-7 was used to identify a human KIM sequence (see AAT96035).  
 CC The isolated KIM nucleic acids can be used in the recombinant  
 CC production of KIM polypeptides in prokaryotic or eukaryotic host  
 CC cells. KIM, or an agonist, can used to treat renal disease and to  
 CC promote the growth of new tissue or the survival of damaged tissue,  
 CC generally in conditions where the binding of specific ligand to KIM  
 CC stimulates cell growth, maintains cellular differentiation or  
 CC reduces apoptosis, e.g. in cases of renal failure, nephritis,  
 CC kidney transplants, toxic or hypoxic injury. Damage/regeneration  
 CC of renal cells can be determined by measuring KIM polypeptide or  
 CC nucleic acid, e.g. with an antisense probe, particularly to  
 CC diagnose or monitor the progress of disease or therapy.  
 XX  
 SQ Sequence 2303 BP; 570 A; 602 C; 578 G; 553 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1, 41e-222 Length: 2303  
 Score: 3040.00 Matches: 568  
 Percent Similarity: 99.83% Conservative: 3  
 Best Local Similarity: 99.30% Mismatches: 1  
 Query Match: 99.31% Indels: 0  
 DB: 19 Gaps: 0  
 US-09-943-075A-2 (1-572) x AAT96034 (1-2303)  
 QY 1 MetCluserteucysglyvalleuValPheLeuLeuLeuAaAGlyLeuProLeuGln 20  
 Db 107 ATGGAAAGTCTGCGGGGCTCGTGTATTTCTGCTGCTGGCTGGAGACTGCGGCTCCAG 166  
 QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
 Db 167 GCGGCCAAGCGTTCGCTGATGCTGGCCATGAGCAGTATCGGATCATCATGAGGGAG 226  
 QY 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro 60  
 Db 227 AACAAACAATACGTGGCTGGCTTCAGATGAAATGAATGGGATGAACAGCTGTATCCA 286  
 QY 61 ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla 80  
 Db 1367 GCCACTCCCAACGGAAGCCTGTACGATCATCTCTGACCCACCTGTCAGATGCCAGAAC 1426

Db 287 GTGTGGAGGAGGGAGAGGCGAGATGGAAGGACTCTCTGGGAAGGAGGCGGTGTCAGGCA 346  
 QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
 Db 347 GCCCTAACCAAGTATTCACCGGCTTGGTGGTTCATATATCACCTTCGTAGTGAACCTG 406  
 QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
 Db 407 GTGTTCCCCAGATGCCAGAAAGATGCCAACGGCAATATCGTCTATGAGAGGAAGTGC 466  
 QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140  
 Db 467 AGAAGTATTTGGAGCTGGCTTCTGACCGGTATGCTACAACCTGGACCACAGGGGAGAC 526  
 QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
 Db 527 GATGAGGACTGGGAAGACACACACAGCCAAAGCCAGCACCTCAGGTTCCCGGAGGAAG 586  
 QY 161 PropPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
 Db 587 CCCTTCCTCGCCCCACGAGCGAAGAAATGGAACCTTCGCTAGCTCTCCACACACTT 646  
 QY 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaAArgValSerIleAsnThrValAsn 200  
 Db 647 GGTCACTATTTTCAAAAGCTGGGTGCTGTCAGCACGAGTTTCTATAAACACACAGTCAAC 706  
 QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220  
 Db 707 TTGACAGTTGGCCCTCAGGTCATGGAGTGATGTTCTTCGAAGACACAGCGCGGCGATAC 766  
 QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
 Db 767 ATTCCCATCTCCAAAGTGAAGACGCTGATGATGATAACAGATCAGATCCCTATATTGGTG 826  
 QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
 Db 827 ACCATGATCCAGAAAGATGACCGGAACCTCGTGTGATGAACCTTCCTCAGACACTCCCC 886  
 QY 261 IlePhePheAspValLeuLeuHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
 Db 887 ATTTTCTTCGATCTCATTCACGATCCAGTCCATTTCTCAACTACTCTGCACTTTCC 946  
 QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
 Db 947 TACAAGTGGAACTTTGGGACACACACTGGCCCTGTTGTTCTCCACACATCAGACTTTGAAT 1006  
 QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
 Db 1007 CACACGTATGTGCTCAATGGAACCTTCAACTTTAACTCAGCTGCAAACTGCAAGTCCG 1066  
 QY 321 GlyProCysProSerProThrProSerProSerSerSerThrSerProSerProAlaSer 340  
 Db 1067 GGACATGCCCCCTCACCCACACTTGGCCCTCTTCTTCGACTTTCCTTCGCTGCACTCT 1126  
 QY 341 SerProSerProThrLeuSerThrProSerProSerProSerLeuMetProThrGlyTyrLysSer 360  
 Db 1127 TCGCCTTACCCACATTTATCAACACTAGTCCCTCTTTAATGCTACTGCGCACAAATCC 1186  
 QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380  
 Db 1187 ATGAGCTGAGTGACATTTCCAAATGAACACTGCCAAATAAACAGATATGGTTACTTCAGA 1246  
 QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400  
 Db 1247 GCCACCATCAATTCATAGATGGAATCCCTAGAAATGCAACATCATCCAGTAGCAGATGTC 1306  
 QY 401 ProfileProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420  
 Db 1307 CCAATCCCAACACCGCGCTGACAACTCAGTATGAGACTTCATTTGACCTGCGAAGGG 1366  
 QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440  
 Db 1367 GCCACTCCCAACGGAAGCCTGTACGATCATCTCTGACCCACCTGTCAGATGCCAGAAC 1426

QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460  
 Db 1427 AGGCTGTGCAGCCGGTGGATGAGCTGTGCTCCTGTCGTCGAGAGACCTTC 1486  
 QY 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeuAlaLeu 480  
 Db 1487 AATGGTCCGGCAGCTACTGTGTGAATTTCACTCTGGGAGAGCATGCAAGCCTGGCCCTC 1546  
 QY 481 ThrSerAlaLeuLeuSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500  
 Db 1547 ACCAGCCCTGATCTATATCCCTGGCAAGACCTAGGCTCCCTCTGAGAACAGTGAAT 1606  
 QY 501 GlyValLeuLeuSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
 Db 1607 GGTCTCTGATCTCCATTTGGCTGCTGGCCATGTTTGTCCACATGTTTACCATCTTCTG 1666  
 QY 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540  
 Db 1667 TCAAAAACACACAGACGTACAAGCCANTAGGAACCTGCACGAGAACGTGGTCAAGGGC 1726  
 QY 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgLysArgGlu 560  
 Db 1727 AAAGCCCTGAGTGTCTTCTCAGCCATGCAAAAGCCCGCTTCTCCGAGGAGACCGGAG 1786  
 QY 561 LysAspProLeuLeuGlnAspLysProTyrMetLeu 572  
 Db 1787 ANGATCCACTGCTCCAGGACCAAGCCATGGATGCTC 1822

RESULT 2  
 AAA96729  
 ID AAA96729 standard; DNA; 2213 BP.  
 AC AAA96729;  
 XX  
 XX  
 DT 19-FEB-2001 (first entry)  
 XX  
 DE Polynucleotide isolated from lymph node stromal cells of fan -/- mice.  
 XX  
 KW Lymph node stromal cell; fan -/- mice; inflammatory disorder;  
 KW immune system disorder; cancer; viral infection; HIV infection;  
 KW blood vessel growth; tumour necrosis factor disorder; arthritis;  
 KW inflammatory bowel disease; fibroblast growth factor-mediated disorder;  
 KW cardiac failure; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 23..1747  
 FT /\*tag= a  
 XX  
 PN WO200058463-A1.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 18-FEB-2000; 2000WO-NZ00015.  
 XX  
 PR 25-MAR-1999; 99US-0276268.  
 PR 26-AUG-1999; 99US-0383586.  
 XX  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;  
 XX Murison JG;  
 DR WPI: 2000-664924/64.  
 DR P-PSDB: AAB19119.  
 XX  
 XX Polypeptide expressed in mammalian fan -/- lymph node stromal cells,  
 PT useful for modulating growth of blood cells, for treating inflammatory  
 PT and tumour necrosis factor-mediated disorders, cancer and viral  
 PT disorders .  
 XX

PS Claim 4; Page 48-49; 75pp; English.

XX The present sequence represents a polynucleotide sequence which is  
 CC isolated from lymph node stromal cells of fan -/- mice. The  
 CC polynucleotides and their polypeptides are useful for treating an  
 CC inflammatory disorder, disorder of immune system and cancer selected  
 CC from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a  
 CC viral disorder, in particular HIV infection and for modulating the  
 CC growth of blood vessels. The polypeptides are useful for treating a  
 CC tumour necrosis factor (TNF) mediated disorder, such as those selected  
 CC from arthritis, inflammatory bowel disease and cardiac failure and a  
 CC fibroblast growth factor-mediated disorder. It is also useful in assays  
 CC to determine biological activity, to raise antibodies, to isolate  
 CC cognate corresponding ligand or receptors, to quantify levels of protein or  
 CC and in compositions for the treatment of skin, connective tissue and  
 CC immune system diseases. The polynucleotide is useful as marker for  
 CC tissue, as a chromosome marker or tags in the identification of a  
 CC genetic disorder.  
 XX

SQ Sequence 2213 BP; 526 A; 599 C; 555 G; 533 T; 0 other;

#### Alignment Scores:

Pred. No.:	1-24e-196	Length:	2213
Score:	2700.00	Matches:	507
Percent Similarity:	92.29%	Conservative:	20
Best Local Similarity:	88.79%	Mismatches:	42
Query Match:	88.21%	Indels:	2
DB:	21	Gaps:	1

US-09-943-075A-2 (1-572) x AAA96729 (1-2213)

QY	1	MetGluSerLeuCysGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln	20
Db	23	ATGGAAGTCTCTCGGGGCTCTGGGATTTCTCTGCTGCTGAGGACTGCTCTCCAG	82
QY	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu	40
Db	83	GCTGCCAAGCATTTCTGATGTCTGGGCATGAACAGTATCCCATCAGAGAG	142
QY	41	AsnAsnGlnLeuArgGlyTyrPsrSerAspGluAsnGluTrpAspGluGlnLeuTyrPro	60
Db	143	CACACCAANTTACGTGGCTGGTCTCGGATGAAATGAATGGATGAACACCTGTATCCA	202
QY	61	ValTrpArgArgGlyGluGlyArgTyrPlyAspSerTrpGluGlyArgValGlnAla	80
Db	203	GTGTGGAGAGGGAGAGCGGAGGCTGGAAGGACTCTCTGGAGAGGCGCTGTGCAGGCA	262
QY	81	AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
Db	263	GTCTGACCACTGACTCACCCTGCTGCTGGTTCCTCAATATACCTTTGTGTGAACCTG	322
QY	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys	120
Db	323	GTGTTCCCCAGATCCAGAGAGAGATGCTTAATGGCAATATCGTCTATGAGAAGAACTGC	382
QY	121	ArgSerAspLeuLeuAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAsp	140
Db	383	AGGAATGATTTGGGACTGACCTCTGACCTGCATGCTTACAACTGGACTGCAGGGCAGAT	442
QY	141	AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys	160
Db	443	GATGTGACTGGGAAGATGGCACCAGCCAGGACGATCTCAGGTTCCTCCGACAGAGAGG	502
QY	161	ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu	180
Db	503	CCCTTCCTCCGCCCTCCATGGATGGAAGAAATAGAGCTTTGTCTACCTCTTTCACACACTT	562
QY	181	GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn	200
Db	563	GGCCAGTATTTCCAAAAAATCGGTGGTGTTCAGCAGCGGGTTTCTATAAACACAGTCAAC	622
QY	201	LeuTyrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr	220

D	b	623	TTGACAGCTGGCCCTCAGGTTCATGGGAAGTGACTGCTTTTTCGAAGATACCAGCGCGGCATAC	682
Q	y	221	IleProIleSerLysValLysAspValTyrrValIleThrAspGlnilePheVal	240
D	b	683	ATTCCCATCTCGAAGGTGAAGATGTGTATGTATAACAGATCAGATCCCTGTATTGCTG	742
Q	y	241	ThrMetTyrrGlnLysAsnAspArgasnSerSerAspGluThrPheLeuArgaspLeupro	260
D	b	743	ACCATGTCCCAGAAGATGACAGGAACATTGCTCATGAGATCTTCCTCAGAGACCTGCC	802
Q	y	261	IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrrSerAlaIleSer	280
D	b	803	ATCGTCTTCGATGTCTTCATTATCATCCAGGCACCTTCCTCAACGACTCTCGCCATTTC	862
Q	y	281	TyrLysTrpAsnPheGlyAspAsnthrGlyLeuPheValSerAsnAsnHisThrLeuAsn	300
D	b	863	TACAAGTGGAACTTTGGGGACAACACTTGGCGTTGTTGTCTCCAACAAATCACACTTTGA	922
Q	y	301	HisThrTyrrValLeuAsnGlyThrPheAsnPheAsnLeuthrValGlnThrAlaValpro	320
D	b	923	CACACTTATGTGCTCAATGGAACCTTAACGCTTAACGCTCACCGCTGCAAACTCGAGTGC	982
Q	y	321	GlyProCysProSerProThrProSer-----ProSerSerSerThrSerProSerPro	338
D	b	983	GGGCCATGCCCTCCCCCTTCGCGCTTCGACTCGCGCTCCACTTCACCTCCGCGCTCAC	1042
Q	y	339	AlaSerSerProSerProThrLeuSerThrProSerProSerLeuMetProthrGlytyr	358
D	b	1043	CGCCCTCAACCTCTGCCACATTATCAACACTAGCCCGCTCTTAATGCTACTGGTTAC	1102
Q	y	359	LysSerMetGluLeuSerAspIleSerAsnGluAsnCysargileAsnArgtyrGlytyr	378
D	b	1103	AAATCCATGGAGCTGAGTGACATTTCCTCAAGAAACTGCCGAATAAACAGATGGCTAC	1162
Q	y	379	PheArgAlaThrIleThrIleValAspGlyIleLeuGluValIleAsnIleIleGlnValala	398
D	b	1163	TTCAGAGCCACCATCACAAATTGTAGAGGGATCCTGGAAGTCAGCATCATGCAGATAGCA	1222
Q	y	399	AspValProIleProthrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCys	418
D	b	1223	GATGTCCCCATGCCACCCAGCCAGCTGCCAACTCCTGATGAGACTCTACTGTGACCTGC	1282
Q	y	419	LysGlyAlaThrProthrGluAluCysthrIleIleSerAspProthrCysGlnIleAla	438
D	b	1283	AAAGSGGCCACCCCATGTGAAGCTGTACGATCATCTCCGACCCCACTGCCAGATGCC	1342
Q	y	439	GlnAsnArgValCysSerProValAlaValAspGluLeucysLeuLeuSerValArgArg	458
D	b	1343	CAGAACCGGCTCTCACCCCTGTGGCTGTGGATGGGCTGTGCCTGCTGTGTGAGAAGA	1402
Q	y	459	AlaPheAsnGlySerGlyThrTyrrCysValAsnPheThrLeuGlyAspAspIleSerLeu	478
D	b	1403	GCCTTCATGGGTCTGCGACCTACTGTGTGAATTTCCTACTCTGGAGATGATGCAAGCTG	1462
Q	y	479	AlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThr	498
D	b	1463	GCCTCACCAGCACCTGATCTCTATCCCTGSCAAGACCCAGACTCCCTCTGAGAGCA	1522
Q	y	499	ValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIle	518
D	b	1523	GTGAATGTGTCTGTATCTCCATTCCTGGCTGCTGTGCTTCTCACCATGTTTACCATC	1582
Q	y	519	LeuLeuTyrrLysLysHisLysThrTyrrLysProIleGlyAsnCysThrArgAsnValVal	538
D	b	1583	TTGCTGTACAAAACACAGCGGTACAGCCCAATAGGAACCTGCCCAAGGAACACGCTC	1642
Q	y	539	LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaproPheSerArgClyAsp	558
D	b	1643	AAGGCCAAAGCCCTGAGTGTCTTCCTACGCCACGCAAGCCCGCTTCTTCGAGGAGAC	1702
Q	y	559	ArgGluLysAspProLeuLeuGlnAspLyspro	569

Db	1703	CAGGAGAAGATCCATTGCTCAGGACAAGCCA	1735
RESULT 3			
AAT69318			
ID	AAT69318	standard; cDNA: 2669 BP.	
XX	AC	AAT69318;	
XX	AC		
XX	AC		
DT	26-FEB-1998	(first entry)	
DE		Murine metastatic nucleic acid sequence.	
XX			
XX		Mouse; murine; tumour; cancer; metastatic sequence; detection;	
KW		diagnosis; treatment; metastasis; hyperplasia; dysplasia;	
KW		hypertrophy; screening; ss.	
XX			
OS		Mus musculus.	
XX			
FN	W09718454-A2.		
XX			
PD	22-MAY-1997.		
XX			
PF	15-NOV-1996;	96WO-U81567.	
XX			
PR	30-JAN-1996;	96US-0594031.	
PR	16-NOV-1995;	95US-0006838.	
XX			
PA	(THOM/) THOMPSON T.		
XX			
PI	Thompson T;		
XX			
XX			
DR	WPI; 1997-289397/26.		
XX			
PT	Identifying tumour metastatic sequences - by introducing transfected cells into host mammal and analysing primary and metastatic sequences by differential display PCR		
PT			
XX			
PS	Disclosure; Fig 12BT; 102pp; English.		
XX			
CC	Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mouse embryos. The UGS cells were infected with retroviruses, cultured and implanted under the renal capsule of mice.		
CC	Reconstitutions were harvested 5 weeks later, when they showed signs of distress from the tumour burden. Metastatised tumours were isolated from a site outside the renal capsule. RNA was isolated from primary tumours and metastases, reverse transcribed and subjected to differential display PCR. The sequences were analysed to obtain metastatic sequences, e.g. the present sequence. The method can be used to detect, diagnose and treat disorders related to metastasis, or treat malignant or non-malignant disorders, e.g. hyperplasia, dysplasia and hypertrophy. The metastatic sequence can be used to screen a biological sample for metastasis, and it or its expression product may also be used to treat a metastatic disorder.		
XX			
SQ	Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 other;		
Alignment Scores:			
Pred. No.:	1-61e-155	Length:	2669
Score:	2163.00	Matches:	394
Percent Similarity:	82.57%	Conservative:	75
Best Local Similarity:	89.37%	Mismatches:	85
Query Match:	70.66%	Indels:	14
DB:	18	Gaps:	1
US-09-943-075A-2 (1-572) x AAT69318 (1-2669)			
Qy	1	MetCluSerLeuCyGlyValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGln	20
Db	92	ATGGAAATGCTCTACTATTTCTGGGATTTCTGCTCTGCTCAGAGATGCCATTCGAT	151
Qy	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluClnPyrProAspHisMetArgGlu	40
Db	152	CGCGCCCAACAGATTCATCATGTGCTGGGCAATGAAGACCTTCTGCTTACATGAGGAG	21

Qy	401	ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly	420
Db	1250	CTGATGCGCGGCGCATCGGCTGAAAGCTCCCTTAATAGACTTCTCGTGACCTGCCAAGGG	1309
Qy	421	AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn	440
Db	1310	AGCATTCCTCCCGAGGTCGTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAAC	1369
Qy	441	ArgValCysSerProValAlaValAspGluLeuCysLeuSerValargargAlaPhe	460
Db	1370	ACAGTCTCGACGCCCTGGGATGGGATGAGATGGTGTCTGTGACTGTGAGAGCAACCTTC	1429
Qy	461	AsnGlySerGlyThrTyrcysValAsnPheThrLeuGlyAspAlaSerLeuAlaLeu	480
Db	1430	AATGGGTCTGGAGCTACTGTGTGAACCTCACCTTGGGGATGACACAAAGCCTGCCTCTC	1489
Qy	481	ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn	500
Db	1490	ACGAGCACCTTGATTTCTGTCTCTGACAGAGACCCGCGCTTAAAGGATGGCAAAC	1549
Qy	501	GlyValIleIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu	520
Db	1550	AGTGCCCTGATCTCCGTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTG	1609
Qy	521	TyrLysLysHisLysThrTyrcysProIleGlyAsnCysThrArgAsnValIleValLysGly	540
Db	1610	TCAAAACACACAGGAAATCAACCCCATAGAAAATAGTCTCTGGGATGTGTCAAGACC	1669
Qy	541	LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu	560
Db	1670	AAA3GCTGAGTGCTTCTTCAACCGTGCAAAAGCCGTGTCTTCCGGGAAACAGGAA	1739
Qy	561	LysAspProLeuLeuGlnAspLys	568
Db	1730	AAGATCGCGCTACTCAAAAACCAA	1753
RESULT 4			
AAT69328			
XX	ID	AAT69328 standard; cdna; 2669 bp.	
XX	AC	AAT69328:	
XX	DT	26-FEB-1998 (first entry)	
XX	DE	Murine metastatic nucleic acid sequence.	
XX	KW	Mouse; murine; tumour; cancer; metastatic sequence; detection;	
XX	KW	diagnosis; treatment; metastasis; hyperplasia; dysplasia;	
XX	KW	hypertrophy; screening; ss.	
XX	OS	Mus musculus.	
XX	PN	W09718454-A2.	
XX	PD	22-MAY-1997.	
XX	PF	15-NOV-1996; 96WO-U81567.	
XX	PR	30-JAN-1996; 96US-0594031.	
XX	PR	16-NOV-1995; 95US-0006838.	
XX	TH	(THOM/)	
XX	TH	THOMPSON T.	
XX	PI	Thompson T;	
XX	DR	WPI; 1997-289397/26.	
XX	PT	Identifying tumour metastatic sequences - by introducing transfected	
XX	PT	cells into host mammal and analysing primary and metastatic	
XX	PS	sequences by differential display PCR	
XX	PS	Disclosure; Fig 12CS; 102pp; English.	

CC Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old  
CC mouse embryos. The UGS cells were infected with retroviruses,  
CC cultured and implanted under the renal capsule of mice.  
CC Reconstitutions were harvested 5 weeks later, when they showed  
CC signs of distress from the tumour burden. Metastasised tumours were  
CC isolated from a site outside the renal capsule. RNA was isolated  
CC from primary tumours and metastases, reverse transcribed and  
CC subjected to differential display PCR. The sequences were analysed  
CC to obtain metastatic sequences, e.g. the present sequence. The  
CC method can be used to detect, diagnose and treat disorders related  
CC to metastasis, or treat malignant or non-malignant disorders, e.g.  
CC hyperplasia, dysplasia and hypertrophy. The metastatic sequence can  
CC be used to screen a biological sample for metastasis, and it or its  
CC expression product may also be used to treat a metastatic disorder.  
xx  
SQ Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 other.

Alignment Scores:				
Pred. No.:	1.61e-155	Length:	2669	
Score:	2163.00	Matches:	394	
Percent Similarity:	82.57%	Conservative:	75	
Best Local Similarity:	69.37%	Mismatches:	85	
Query Match:	70.66%	Indels:	14	
DB:	18	Gaps:	1	
US-09-943-075A-2 (1-572) x AAT69328 (1-2669)				
Qy	1	MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGln	20	
Db	92	ATGGAATGTCCTACTATTCTCGGATTTCCTCTCGCTGCGTCAAGATTGCCACTTGAT	151	
Qy	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu	40	
Db	152	GCGGCCAAACGATTTCATGATGTCTGCTGGGCAATGAAGACCTTCGTCTACATGAGGAG	211	
Qy	41	AsnAsnGlnLeuArgGlyTyrPsrSerAspGluAsnGlnTyrAspGluGlnLeuTyrPro	60	
Db	212	CACATCAATTAATGCGTGTCTTCGTGATGAATGACTGCGAATGAAAACTCTACCCA	271	
Qy	61	ValTyrArgArgGlyGluGlyArgTyrPlysAspSerTyrGluGlyGlyArgValGlnAla	80	
Db	272	GTCTGGAACGGGGAGACATCAGGTGGTGAANAATCTCTGGAAGGAGGCCGTGTGCAGGCG	331	
Qy	81	AlaLeuThrSerAspSerProAlaLeuValValGlySerAsnIleLeuPheValValAsnLeu	100	
Db	332	GTCTGACCAGTGACTCACCAGCCCTCGTGGGCTCAAAATATAACATTTCGCGTGAACCTG	391	
Qy	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys	120	
Db	392	ATATATCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAAGTCG	451	
Qy	121	ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrThrGlyAlaAsp	140	
Db	452	AGAAATGAGCGCTGGTTATCTGCTGATCCATATGTTTACACTGGACAGCATGGTCAGAG	511	
Qy	141	AspGluAspTyrPgluAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys	160	
Db	512	GACAGTGACGGGGAATATGGACCGGCACCAAGCATCATAAACGTCTTCCTGTATGGGAAA	571	
Qy	161	ProPheProArgProHisGlyLysLysTyrAsnPheValTyrValPheHisThrLeu	180	
Db	572	CCTTTTCCTACCACCCCGGATGGAGAAGATGGAATTCATCTACGTACGCTCCACACACTT	631	
Qy	181	GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn	200	
Db	632	GGTCACTATTTCGAAATTTGGGACGATGTTCAGTCAGAGTTTCTGTGAACACAGCCAAAT	691	
Qy	201	LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr	220	
Db	692	GTGACACTTGGCGCTCAACTCATGGAAAGTGACTGTCTACAGAAGACATGGACGGGCATAT	751	
Qy	221	IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal	240	



XX AAC55715;  
AC  
XX  
XX  
XX 17-JAN-2001 (first entry)  
XX  
XX Human NMB cDNA sequence from Genbank X76534.  
DE  
XX  
XX Human; differentially regulated gene; macrophage development;  
KW diagnosis; matrix metalloproteinase 19; MMP19; antiarthritic;  
KW antiinflammatory; destructive macrophage development inhibitor;  
KW arthritis; colorectal cancer; immune response; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200055373-A2.  
PN  
XX  
XX 21-SEP-2000.  
PD  
XX  
XX 15-MAR-2000; 2000WO-US06883.  
PF  
XX  
XX 15-MAR-1999; 99US-0124530.  
PR  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
PA  
XX  
XX Murray R;  
PI  
XX  
XX WPI; 2000-628200/60.  
DR  
XX  
XX  
XX Screening drug candidates comprises adding a drug to a cell expressing  
PT an expression profile gene and determining the effect of the drug on  
PT the expression of the expression profile gene  
XX  
XX Claim 1; Page -: 99pp; English.  
PS  
XX  
XX The present invention describes a method for screening drug candidates.  
CC The method comprises adding a drug to a cell that expresses an  
CC expression profile gene encoding a protein encoded by 5 sequences of  
CC defined base pairs as given in C55638, C55642, C55643, C55644 and  
CC C55653 or a sequence represented by Genbank accession number X92521,  
CC X62466, J04130, X62087 and X76534 (or a fragment) and determining the  
CC effect of the drug on the expression of the expression profile gene.  
CC An inhibitor of matrix metalloproteinase 19 (MMP-19), preferably an  
CC antibody, is useful for treating destructive macrophage disorders (DMD)  
CC by inhibiting DM development in a cell of an individual having  
CC arthritis. Antibodies to MMP-19 are useful for localising a therapeutic  
CC moiety preferably cytotoxic agent or a radioisotope to colorectal cancer  
CC tissue. A composition comprising MMP-19 is useful for eliciting an  
CC immune response in an individual. C55635 to C55710 represent human  
CC differentially regulated genes of the invention. The present sequence  
CC represents the human NMB cDNA sequence according to the Genbank  
CC accession number X76534.  
CC N.B. The present sequence is not given in the present specification,  
CC but it is specifically claimed by its Genbank accession number.  
XX  
SQ Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 other;

Alignment Scores:  
Pred No.: 1 61e-155 Length: 2669  
Score: 2163.00 Matches: 394  
Percent Similarity: 82.57% Conservative: 75  
Best Local Similarity: 69.37% Mismatches: 85  
Query Match: 70.66% Indels: 14  
DB: 21 Gaps: 1

US-09-943-075A-2 (1-572) x AAC55715 (1-2669)

QY 1 MetGluSerLeuGlyValLeuValPheLeuLeuLeuLeuAlaGlyLeuProLeuGln 20  
Db  
92 ATGAATGCTCTACTACTATTTCTGGGATTTCTGCTCGCTGGCTCAAGATTGCCACTTGAT 151  
QY 21 AlaAlaLysArgPheArgValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
Db  
152 GCCCCAAACGATTTCATGATGCTGGGCAATGAAGACCTTCCTGCTTACATGAGGAG 211

QY 41 AsrAsnGlnLeuArgGlyTyrSerSerAspGluAsnGluTyrAspGluGlnLeuTyrPro 60  
Db  
212 CACAATCAATTAAATGGCTGCTCTCTGATGAATAAGTACCTGGAAATGAANAACCTACCCA 271  
QY 61 ValTyrArgGlyGlyGluGlyArgTyrLysAspSerTyrPoloGlyGlyArgValGlnAla 80  
Db  
272 GTGTGAAGCGGGAGACATGAGCTGGAAACCTCTCTGGAAGGAGCGCTGCGAGCG 331  
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
Db  
332 GTCCTGACCACTGACTCACCAGCCCTCTGGGCTCAATATATAACATTTCCGGTGAACTG 391  
QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
Db  
392 ATATTCCCTAGATGCCAAAGAGATGCCAATGGCAACATAGCTATGAGAAGAACTGC 451  
QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrThrGlyAlaAsp 140  
Db  
452 AGAATGAGGCTGGTTTATCTCTGATCCCATATGTTTACAACTGGACACCATGGTCAGAG 511  
QY 141 AspGluAspTyrGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProaspGlyLys 160  
Db  
512 GACAGTGACGGGAAATGGCAGCCGCAAGCCATCATAACTCTTCCCTGATGGAAA 571  
QY 161 ProPheArgProHisGlyArgLysLysTyrPAsnPheValTyrValPheHisThrLeu 180  
Db  
572 CCTTTCTCTACCCCGGATGGAGAAATGGAATTTTATCTACCTGCTTCCACACACTT 631  
QY 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
Db  
632 GGTACGATTTTCCAGAAATGGGACGATGTTGAGTGAGAGTTTCTGTGAACACAGCAAT 691  
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220  
Db  
692 GTGACACTTGGGCTCTCAACTCATGGAAGTGACTGCTTACAGAAAGACATGGAGCGCATAT 751  
QY 221 IlePheLeuSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
Db  
752 GTTCCCATCGCACAAAGTGTCTACGTGGTAACAGATCAGATTCCTGTGTTGTG 811  
QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
Db  
812 ACTATGTTCCAGAGACAGATCGAAATTCATCCGAGAAACCTTCTCAAGATCTCCCC 871  
QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
Db  
872 ATATGTTTGTGCTCTGATTCATGATCCTAGCCACTTCTCAATTTATTTACCATTAAC 931  
QY 281 TyrLysTyrAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
Db  
932 TACAAGTGGAGCTTGGGGGATTAATCTGGCTGTTTGTGTTTCCACCACATCATCTGTGAAT 991  
QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheLeuThrValGlnThrAlaValPro 320  
Db  
992 CACAGTATGTCTCAATGGAACCTTCAGCCTTAACTCTGTAAGAACTGCAGCACCA 1051  
QY 321 GlyProCysProSerProThrProThrProSerSerSerThrSerProSerProAlaSer 340  
Db  
1052 GGAGCTTGTCTCCGCCACCCGCCACCCACCCAGACCT----- 1087  
QY 341 SerProSerProThrLeuSerThrProSerProSerSerSerLeuMetProThrGlyTyrLysSer 360  
Db  
1088 -----TCNAACCCACCCCTTCTTAGGACCTGCTGGTGCAACCCC 1129  
QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380  
Db  
1130 CTGGAGCTGAGTAGGATTCCTGATGAATAACTGCCAGATTAAACAGATATGCCACTTCAA 1189  
QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAlaSer 400  
Db  
1190 GCCACCATCACAAATTTAGAGGGAACTTTAGAGGTTAAACATCATCATCAGATGACAGACGTC 1249

QY 401 ProfileProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420  
Db 1250 CTGATCCGGTGGCCATGGCTGAAGCTCCCTAATAGACTTGTGCTGACCTCCCAAGG 1309  
QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440  
Db 1310 AGCATTTCCACCGAGGTGTGTACCATATTTCTGACCCCACTGCGGAGATCACCCAGAAC 1369  
QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuSerValArgArgAlaPhe 460  
Db 1370 ACAGTCTGCAGCCCTGTGGATGGATGAGATGTCTGCTGACTGTGACGACGAACCTTC 1429  
QY 461 AsnGlySerGlyThrTyrcysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480  
Db 1430 AATGGGTCTGGACCTACTGTGTGAACCTCACCTGGGGGATGACACAGCCCTGGCTCTC 1489  
QY 481 ThrSerAlaLeuIleSerIleProGlyLysLeuGlySerProLeuArgThrValAsn 500  
Db 1490 ACGAGCACCTGATTTCTGCTCTGACAGAGACCCAGCTCGCTTTAAGATGGCAAC 1549  
QY 501 GlyValLeuIleSerIleCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
Db 1550 AGTCCCTGATCTCGTGGCTGTGGCCATTTGTGCTGACTGTGATCTCCCTCTTGGTG 1609  
QY 521 TyrLysLysHisLysThrTyrcysValAsnProIleGlyAsnCysThrArgAsnValLysGly 540  
Db 1610 TACAAAAACACAGGAATACACACCAATAGAAATAGTCTGGGAATGTGGTCAGAACG 1669  
QY 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560  
Db 1670 AAGGCTGAGTGTCTTCTCAACCGTGCNAAGCCGTGTCTTCCCGGGAACACAGGAA 1729  
QY 561 LysAspProLeuGlnAspLys 568  
Db 1730 AAGGATCCGCTACTCAAAAAACAA 1753  
RESULT 6  
ABQ88185  
ID ABQ88185 standard; cDNA; 2669 BP.  
XX AC ABQ88185;  
XX 18-SEP-2002 (first entry)  
XX Human osteoblast differentiation related cDNA SEQ ID NO 92.  
XX Human; osteoblast; stem cell differentiation; bone tissue deposition;  
KW osteoporosis; osteopathic; ss.  
XX Homo sapiens.  
XX WO200250301-A2.  
XX 27-JUN-2002.  
XX 18-DEC-2001; 2001WO-US48276.  
XX 18-DEC-2000; 2000US-255882P.  
XX 24-APR-2001; 2001US-285691P.  
XX (GENE-) GENE LOGIC INC.  
PA (PROC) PROCTER & GAMBLE CO.  
XX Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;  
PI Mertz L;  
XX WPI; 2002-557663/59.  
XX Use of genes and their expression profiles associated with osteoblast  
PT differentiation for screening modulators bone formation, for diagnosing  
PT or treating e.g. osteoporosis, or as markers for the differentiation  
PT process -  
XX

PS Claim 1; SEQ ID NO 92; 78pp + Sequence Listing; English.  
XX The invention relates to genes and their expression profiles are used  
CC for:  
CC (a) screening modulators of precursor stem cell differentiation into  
CC osteoblasts, or bone tissue deposition;  
CC (b) diagnosing abnormal deposition of bone tissue, abnormal rate of  
CC osteoblast formation or osteoporosis; or  
CC (c) treating or monitoring treatment of the conditions cited in (b), or  
CC monitoring the progression of bone tissue deposition.  
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid  
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy  
CC drug-induced abnormalities in bone formation or bone loss, conditions  
CC that involve altered bone metabolism (e.g. idiopathic juvenile  
CC osteoporosis), skeletal disease linked to breast cancer, mastocytosis,  
CC Fanconi syndrome or fibrous dysplasia. The present sequence is that of an  
CC osteoblast differentiation associated cDNA marker of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 other;

#### Alignment Scores:

Pred. No.:	1.61e-155	Length:	2669
Score:	2163.00	Matches:	394
Percent Similarity:	82.57%	Conservative:	75
Best Local Similarity:	69.37%	Mismatches:	85
Query Match:	70.66%	Indels:	14
DB:	24	Gaps:	1

US-09-943-075a-2 (1-572) x ABQ88185 (1-2669)

Qy	1	MetGluSerLeuCysGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln	20
Db	92	ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTCTGCTGGCTGCAAGATTCGCCATTGAT	151
Qy	21	AlaAlaLysArgPheArgaspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu	40
Db	152	GCCGCCCAAGCATTTTCATGTGTGCGGCAATGAAGACCTCTGCTTACATGAGGAG	211
Qy	41	AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro	60
Db	212	CACAATCAATTAATGGCTGGCTTCTGATGAAATGACTGGAATGAAAAAATCTACCCA	271
Qy	61	ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla	80
Db	272	GTGTGGAGGGGGAGCATGAGGTGGAAAAAATCTCTGGAAGGAGGCCGTGTCAGCGC	331
Qy	81	AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
Db	332	GTCTGACCACTGACTCACCAGCCCTCGTGGCTCAATATAAATTTGCGGTGAACCTG	391
Qy	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys	120
Db	392	ATATTCCTTAGATGCCAAAGGAAGATGCCAATGCAATGCAATAGTCTATGAGAAGAAGCTGC	451
Qy	121	ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp	140
Db	452	AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTACAACTGGACAGCATGCTCAGAG	511
Qy	141	AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys	160
Db	512	GACAGTGACGGGAAAAATGGCAGCCGCCAAAGCCCATCAACGCTCTCCCTGATGGGAAA	571
Qy	161	ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu	180
Db	572	CCTTTTCCTCACCAACCCCGGATGGAGAGATGGAATTTCACTAGCTCTCCACACACTT	631
Qy	181	GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn	200
Db	632	GGTCAAGTATTTCAGAAATTTGGAGCATGTTCTCAGTGAGAGTTTCTGTGAACACACCAAT	691

Oy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220  
Db 692 GTGACACTGGCCCTCAACTCATGGAGTGACTGTCTACAGAGACATGGACGGCATAT 751  
Oy 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
Db 752 GTTCCCATCCGACAAAGTGAAGTGTGTACGTGGTAACAGATCAAGATTCCTGTGTTGTG 811  
Oy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
Db 812 ACTATGTTCCAGAAGAACGATCGAATTCATCCGAGAACTTCTCTCAAGATCTCCCC 871  
Oy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
Db 872 ATTATGTTTGTATGCTGATTCATGATCTAGCCACTTCTCTCAATATATTACCATTAAC 931  
Oy 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
Db 932 TACAAGTGGAGCTCGGGGATAATCTGGCCTGTTTGTTCACCAATCATATCTGTGAAT 991  
Oy 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
Db 992 CACACGTATGTGCTCATGGAACCTTCAGCCTTAACCTCACTGTGAAAGTGCACGACCA 1051  
Oy 321 GlyProCysProSerProThrProSerProSerSerSerThrSerProAlaSer 340  
Db 1052 GCACCTTGTGCGCACCGCCACCCACCCACCCAGACCT----- 1087  
Oy 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyrLysSer 360  
Db 1088 -----TCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGCAACACCC 1129  
Oy 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380  
Db 1130 CTGAGCTGATGAGTCTCTGATGAAACTGCCAGATTAAACAGATATGGCCACTTTCAA 1189  
Oy 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400  
Db 1190 GCCACCATCAATTTGTAGAGGAATCTTAGAGTTAAATCATCATCCAGATGACAGAGCTC 1249  
Oy 401 ProfileProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420  
Db 1250 CTGATCGCGTGGCTGCGCTGAAAGCTCCCTAATAGACTTGTGCTGACCTGCCAAGGG 1309  
Oy 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440  
Db 1310 AGCATTCGCCAGGAGTCTGTACCATCATTTCTTGACCCACCCACCTGGGAGATCACCAGAAC 1369  
Oy 441 ArgValCysSerProValAlaValAspGluLeuCysLeuSerValArgAlaPhe 460  
Db 1370 ACAGCTGCGACCCCTGGGATGTGGATGAGATGTCTGCTGACTGTGAGACGACCTTC 1429  
Oy 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480  
Db 1430 AATGGGTCTGGGAGCTACTGTGAACTTCACTGCGGGGATGACACAAAGCTGCTCTC 1489  
Oy 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500  
Db 1490 ACGAGCACCTGATTCTGTCTCTGACAGACCCAGCCCTCCCTTTAGGATGGCAAC 1549  
Oy 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
Db 1550 AGTGCCTGATCTCCGTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCTCTTGGTG 1609  
Oy 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540  
Db 1610 TACAAAAACAAAGGAATCAACCCCAATAGAAATAGTCTGGGAATGTGGTCAGAGC 1669  
Oy 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560  
Db 1670 AAAGGCTGATGTCTTCTCAACCGTCGCAAAAGCGGTCTTCTCCCGGGAACACGAA 1729  
Oy 561 LysAspProLeuLeuGlnAspLys 568

Db 1730 AAGGATCCGCTACTCAAAACCAA 1753  
RESULT 7  
ABK33556  
ID ABK33556 standard; cDNA; 2683 BP.  
XX AC ABK33556;  
XX 08-MAY-2002 (first entry)  
XX cDNA encoding human PRO protein, Seq ID No 41.  
XX Human; secreted protein; PRO; tumour; lung cancer; colon cancer;  
KW breast cancer; prostate tumour; rectal tumour; liver tumour;  
KW pericyte cell proliferation; chondrocyte cell proliferation;  
KW tumour necrosis factor-alpha; gene; ss.  
XX OS Homo sapiens.  
XX PN WO200208288-A2.  
XX PD 31-JAN-2002.  
XX PF 29-JUN-2001; 2001WO-US21066.  
XX PR 20-JUL-2000; 2000US-219556P.  
PR 25-JUL-2000; 2000US-220585P.  
PR 25-JUL-2000; 2000US-220605P.  
PR 25-JUL-2000; 2000US-220607P.  
PR 25-JUL-2000; 2000US-220624P.  
PR 25-JUL-2000; 2000US-220638P.  
PR 25-JUL-2000; 2000US-220664P.  
PR 25-JUL-2000; 2000US-220666P.  
PR 26-JUL-2000; 2000US-220893P.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 15-SEP-2000; 2000US-000000P.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 28-NOV-2000; 2000US-253646P.  
PR 01-DEC-2000; 2000WO-US32878.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 10-MAY-2001; 2001US-0854280.  
PR 25-MAY-2001; 2001WO-US17092.  
XX (GETH ) GENENTECH INC.  
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;  
XX WPI; 2002-172001/22.  
DR P-PSDB; FAU83612.  
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,  
PT useful for treating a PRO related disorder and for diagnosing tumours  
PT such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal  
PT tumour or liver tumour -  
XX Claim 2; Figure 41; 359pp; English.  
XX The invention relates to one hundred and twenty two nucleic acids  
CC encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides  
CC encode human secreted proteins. The PRO nucleic acids, polypeptides,  
CC agonists and antagonists are useful for treating a PRO related disorder.  
CC The PRO polypeptides are useful for diagnosing tumours, especially lung  
CC cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or  
CC liver tumour. The PRO polypeptides are useful for stimulating the  
CC proliferation of, or gene expression, in pericyte cells, for stimulating  
CC the proliferation or differentiation of chondrocyte cells, for  
CC stimulating the release of tumour necrosis factor-alpha from human blood,

CC for stimulating or inhibiting the proliferation of normal human dermal  
 CC fibroblast cells. The PRO polypeptide may also be used as molecular  
 CC weight markers and for tissue typing. The PRO nucleic acids have  
 CC applications in molecular biology, including use as hybridisation probes,  
 CC and in chromosome and gene mapping. ABK33536-ABK33657 represent human  
 CC PRO protein coding sequences of the invention.

XX SQ Sequence 2683 BP; 751 A; 604 C; 596 G; 732 T; 0 other;

Alignment Scores:  
 Pred. No.: 1,93e-155 Length: 2683  
 Score: 2162.00 Matches: 396  
 Percent Similarity: 82.72% Conservatives: 78  
 Best Local Similarity: 59.11% Mismatches: 87  
 Query Match: 70.63% Indels: 12  
 DB: 24 Gaps: 2

US-09-943-075A-2 (1-572) x ABK33556 (1-2683)

QY 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaAGlyLeuProLeuGln 20  
 DB 82 ATGGAAATGCTCTACTATTCTCGGATTTCTGCTCTGCTGCTGCAAGATTGCCACCTTGAT 141  
 QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
 DB 142 GCGGCCAAACGATTTTCATGATGCTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 201  
 QY 41 AsnAsnGlnLeuArgGlyTrpSerAspGluAsnGlnTrpAspGluGlnLeuTyrPro 60  
 DB 202 CACAAATCAATTAATGGCTGCTCTCTGATGCAAAATGACTGGATGAAATCTACCCA 261  
 QY 61 ValTrpArgGlyGlyGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla 80  
 DB 262 GTGTGAAGCGGGAGACATCAGGTGGTGAATACTCTCGAAGGAGGCCGTGTGACAGGG 321  
 QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
 DB 322 GTCTCACCAGCTGACTCACCACCCCTCTGGGCTCAATATATACATTGGCGTGAACCTG 381  
 QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
 DB 382 ATATTCCTCCATAGATGCAAAAGCAAGATGCCAATGCCAATAGTCTATGAGAAGAACTGC 441  
 QY 121 ArgSerAspGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAsp 140  
 DB 442 AGAAATGAGGCTGGTTATCTGCTGATCCGTATGTTTACAACTGCACAGCATGCTCAGAG 501  
 QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
 DB 502 GACAGTGACGGGGGAAATGGCACCAGCAAGCCATCATAACGCTCTTCCCTGATGGGAAA 561  
 QY 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
 DB 562 CTTTTTCTCACACCCCGGATGGAGAGATGGAAATTCATCTACGCTCTCCACACACTT 621  
 QY 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
 DB 622 GGTCAATATTCAGAAATTTGGACGATGTTCTAGTGAGAGTTTCTGTAACACAGCAAT 681  
 QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220  
 DB 682 GTGACACTTGGGCTCACTCATGGAAGTGACTGTCTACAGAACATGACCGGGCATAT 741  
 QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIlePheVal 240  
 DB 742 GTTCCCATCGCAAGTGAAGATGTGACGTGTTAATCAGATCAGATTCCTGTTGTTGTG 801  
 QY 241 ThrMetTyrClnLysAsnAspArgAsnSerSerAspGlnThrPheLeuArgAspLeuPro 260  
 DB 802 ACTATGTTCCAGAAACCATGAAATTCATCCGACCAAACTTCTTCAAGATCTCCCC 861  
 QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280

DB 862 ATTATGTTTATGCTCTCTGATTCATGATCCTAGCCACTTCTCTCAATATTCTTACATTAAC 921  
 QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn 300  
 DB 922 TACAAGTGGAGCTTCGGGATAATACTGCGCTGTTGTTTCCACCAATCATACTGTGAAT 981  
 QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
 DB 982 CACACGTATGTCTCAATGAACTTCAGCTTAACTCACTGAAAGCTGCAAGACCA 1041  
 QY 321 GlyProCysProSerProThrProSerProSerSerThrSerProSerProAlaSer 340  
 DB 1042 GGACCTGTCCGGCACCGCCACCACCACCCAGA-----CCITCA 1080  
 QY 341 SerProSerProThrLeuSerThr-----ProSerProSerLeuMetPro 355  
 DB 1081 AAACCCACCCCTTCTTTAGCACTCTCTAAATCTTATGATTCAACACCCCGAGGACCT 1140  
 QY 356 ThrGlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnGlyCysArgIleAsnArg 375  
 DB 1141 ACTGGTGACAAACCCCTGGAGCTGAGTAGGATCTCTGATGAAACTGCCAGATTAAACAGA 1200  
 QY 376 TyrGlyTyrPheArgAlaThrIleThrIleValaspGlyIleLeuGluValAsnIleIle 395  
 DB 1201 TATGGCCACTTTCAAGCCACCACATCAATTTAGAGGAATCTTAGAGGTTAAACATCATC 1260  
 QY 396 GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415  
 DB 1261 CAGATGACAGACGCTCTGATGCGGTGCCCTGGCTGAAAGCTCCCTAATAGACTTTGTC 1320  
 QY 416 ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435  
 DB 1321 GTGACCTGCCAAGGGAGCATTCCTCCAGGAGGCTCTGATCATTCTTGACCCACCTGTC 1380  
 QY 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValaspGluLeuCysLeuLeuSer 455  
 DB 1381 GAGATCACCACAGAACACAGCTGCCAGCCCTGTGGATGGATGAGATGCTGCTGCTGACT 1440  
 QY 456 ValArgArgAlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAsp 475  
 DB 1441 GTGAGACGAACTTCAATGGGTGGGAGCTACTGTGAACTCACCCTGGGGGATGAC 1500  
 QY 476 AlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerPro 495  
 DB 1501 ACAAGCTGCTCTCAGGACACCTCTGATTTCTGTTCTCTGACAGAGACCCAGCTCGCT 1560  
 QY 496 LeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMet 515  
 DB 1561 TTAAGGATGGCAACACAGTCCCTGATCTCCGTGGCTGCTTGGCCATATTGTCACGTGTG 1620  
 QY 516 ValThrIleLeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArg 535  
 DB 1621 ATCTCCCTCTTGGTGTACAAAACACAGGAATACACCCCAATAGAAATAGTCTGGG 1680  
 QY 536 AsnValValLysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSer 555  
 DB 1681 AATGTGTCAGAAAGCAAGGCTGAGTGTCTTCTCAACCGTCGCAAAAGCGCTGTTCTTC 1740  
 QY 556 ArgGlyAspArgGluLysAspProLeuLeuGlnAspLys 568  
 DB 1741 CCGGAAACACAGAAAGGATCGCTACTCAAAACCA 1779

RESULT 8  
 AATG9340  
 ID AATG9340 standard; cDNA; 620 BP.  
 XX AC AATG9340;  
 XX XX 26-FEB-1998 (first entry)  
 DE Murine metastatic nucleic acid sequence.  
 XX Mouse; murine; tumour; cancer; metastatic sequence; detection;  
 KW

KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;  
 KW hypertrophy; screening; ss.  
 XX Mus musculus.  
 XX WO9718454-A2.  
 XX PN 22-MAY-1997.  
 XX PD 15-NOV-1996; 96WO-US18567.  
 XX PF 30-JAN-1996; 96US-0594031.  
 XX PR 16-NOV-1995; 95US-0006838.  
 XX PA (THOM/) THOMPSON T.  
 XX PI Thompson T;  
 XX DR WPI: 1997-289397/26.  
 XX PT Identifying tumour metastatic sequences - by introducing transfected  
 XX cells into host mammal and analysing primary and metastatic  
 XX sequences by differential display PCR  
 XX PS Disclosure; Fig 12DD; 102pp; English.  
 XX CC Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old  
 CC mouse embryos. The UGS cells were infected with retroviruses,  
 CC cultured and implanted under the renal capsule of mice.  
 CC Reconstructions were harvested 5 weeks later, when they showed  
 CC signs of distress from the tumour burden. Metastatised tumours were  
 CC isolated from a site outside the renal capsule. RNA was isolated  
 CC from primary tumours and metastases, reverse transcribed and  
 CC subjected to differential display PCR. The sequences were analysed  
 CC to obtain metastatic sequences, e.g. the present sequence. The  
 CC method can be used to detect, diagnose and treat disorders related  
 CC to metastasis, or treat malignant or non-malignant disorders, e.g.  
 CC hyperplasia, dysplasia and hypertrophy. The metastatic sequence can  
 CC be used to screen a biological sample for metastasis, and it or its  
 CC expression product may also be used to treat a metastatic disorder.  
 XX SQ Sequence 620 BP; 143 A; 207 C; 128 G; 142 T; 0 other;

Alignment Scores:  
 Pred. No.: 5.05e-63 Length: 620  
 Score: 941.00 Matches: 183  
 Percent Similarity: 91.75% Conservative: 6  
 Best Local Similarity: 88.83% Mismatches: 15  
 Query Match: 30.74% Indels: 3  
 DB: 18 Gaps: 1

US-09-943-075A-2 (1-572) x AAT69340 (1-620)

QY 274 LeuAsnTyrSerAlaIleSerTyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheVal 293  
 DB 2 CTCACGACTGTGCATTTCTCAAGTGGAACTTTGGGGACAACACTGGCTGTGTC 61  
 QY 294 SerAsnAsnHlsThrLeuAsnHlsThrTyrValLeuAsnGlyThrPheAsnPheAsnLeu 313  
 DB 62 TCCACACATCACACTTTGATATCACACTTATGTGCTCAATGGAACTTCAACCTTAACCTC 121  
 QY 314 ThrValGlnThrAlaValProGlyProCysProSerProThrProSerProSerProSer 331  
 DB 122 ACCGTGCAAACTGAGTGGCCGGCCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCA 181  
 QY 332 SerSerThrSerProSerProAlaSerSerProSerProThrLeuSerThrProSerPro 351  
 DB 182 CTTCAACTCGCCCTCACCTCCGCCCTCACCTCTGCCACATTTATCAACACTAGCCCC 241  
 QY 352 SerLeuMetProThrGlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnCys 371  
 DB 242 TCTTTAATGCCTAGTGTTCATAATCATTCATGGAGCTGAGTGACATTTCCAAATGAACCTGC 301

QY 372 ArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGlu 391  
 DB 302 CGAATAACAGATATGGCTACTTCAGAGCCACCACCATCACAATTTAGAGGGGATCCTGGAA 361  
 QY 392 ValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeu 411  
 DB 362 GTCAGCATCATGAGATAGCAGATGTCCCATGCCACACCCAGCCTGCCAACTCCCTG 421  
 QY 412 MetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSer 431  
 DB 422 ATGACTTCTACTGTGACCTGCAAGGGGCCACCCCATGGAAGCCTGTAGAGTCATCTCC 481  
 QY 432 AspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeu 451  
 DB 482 GACGCCACCTGCCAGATGCCCCAGAACCGGGTCTGCAGCCCTGTGGTGTGGATGGGCTG 541  
 QY 452 CysLeuSerValArgAlaPheAsnGlySerGlyThrTyrCysValAsnPheThr 471  
 DB 542 TGGTG-CTGTCTGTGAGAGAGCCTTCAATGGGTCTGGACCTACTGTGTGAATTTCACT 600  
 QY 472 LeuGlyAspAspAlaSer 477  
 DB 601 CTGGAGATGATGCAAGC 618

RESULT 9  
 AAT69332/C  
 ID AAT69332 standard; cDNA: 524 BP.  
 XX AC AAT69332;  
 XX DT 26-FEB-1998 (first entry)  
 XX DE Murine metastatic nucleic acid sequence.  
 XX KW Mouse; murine; tumour; cancer; metastatic sequence; detection;  
 KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;  
 KW hypertrophy; screening; ss.  
 XX OS Mus musculus.  
 XX PN WO9718454-A2.  
 XX PD 22-MAY-1997.  
 XX PF 15-NOV-1996; 96WO-US18567.  
 XX PR 30-JAN-1996; 96US-0594031.  
 XX PR 16-NOV-1995; 95US-0006838.  
 XX PA (THOM/) THOMPSON T.  
 XX PI Thompson T;  
 XX DR WPI: 1997-289397/26.  
 XX PT Identifying tumour metastatic sequences - by introducing transfected  
 XX cells into host mammal and analysing primary and metastatic  
 XX sequences by differential display PCR  
 XX PS Disclosure; Fig 12CV; 102pp; English.  
 XX CC Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old  
 CC mouse embryos. The UGS cells were infected with retroviruses,  
 CC cultured and implanted under the renal capsule of mice.  
 CC Reconstructions were harvested 5 weeks later, when they showed  
 CC signs of distress from the tumour burden. Metastatised tumours were  
 CC isolated from a site outside the renal capsule. RNA was isolated  
 CC from primary tumours and metastases, reverse transcribed and  
 CC subjected to differential display PCR. The sequences were analysed  
 CC to obtain metastatic sequences, e.g. the present sequence. The  
 CC method can be used to detect, diagnose and treat disorders related  
 CC to metastasis, or treat malignant or non-malignant disorders, e.g.  
 CC hyperplasia, dysplasia and hypertrophy. The metastatic sequence can

CC be used to screen a biological sample for metastasis, and it or its  
 CC expression product may also be used to treat a metastatic disorder.  
 XX  
 SQ Sequence 524 BP; 106 A; 136 C; 163 G; 119 T; 0 other;

Alignment Scores:  
 Pred. No.: 6.29e-47 Length: 524  
 Score: 729.00 Matches: 150  
 Percent Similarity: 89.66% Conservative: 6  
 Best Local Similarity: 86.21% Mismatches: 18  
 Query Match: 23.82% Indels: 2  
 DB: 18 Gaps: 0

US-09-943-075A-2 (1-572) x AAT69332 (1-524)

QY 389 IleLeuGluValAsnIleGlnValAlaAspValProThrLeuGlnProAsp 408  
 DB 522 ATCTGGAAGTCAGCATCATGACATGATGCTCCCATGCCACACCGGAGCTGCC 463  
 QY 409 AsnSerLeuMetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCysThr 428  
 DB 462 AACTCCCTGATGACTTCTGCTGCTGCAAGAGGGCCACCCCATGGAAGCCTGTACG 403  
 QY 429 IleIleSerAspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAlaVal 448  
 DB 402 ATCATCTCCGACCCACCTGCCAGATGCCAGAGAACCGGGCTGTCAGCCCTGTGGCTGTG 343  
 QY 449 AspGluLeuCysLeuLeuSerValArgAlaPheAsnGlySerGlyThrThrCysVal 468  
 DB 342 GATGGGCTGTGTG-CTGTCTGTGAGAGAGCCCTCAATGGGTCTGGCACCCTACTGTGTG 284  
 QY 469 AsnPheThrLeuGlyAspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIlePro 488  
 DB 283 AATTTCACTCTGGGAGATGATGAGCTGGCCCTCACACGACCTGATCTATATCCCT 224  
 QY 489 GlyLysAspLeuGlySerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCys 508  
 DB 223 GCAAGAGACC-AGACTCCCTCTGAGAGCAGTGAATGGTGTCTGATCTCCATCGGCTGC 165  
 QY 509 LeuAlaMetPheValThrMetValThrIleLeuLeuThrLysLysLysThrThrLys 528  
 DB 164 CTGGCTGTGCTGTGTCACCATGTTACCATCTTGTGTACAAAAACACAGGCGTACAG 105  
 QY 529 ProIleGlyAspCysThrArgAsnValValLysGlyLysGlyLeuSerValPheLeuSer 548  
 DB 104 CCAATAGGAACCTGCCCGAGGACACCGTCAAGGGCAAGGGCTGAGTGTCTCTCACT 45  
 QY 549 HisAlaLysAlaProPheSerArgLysAspArgGluLysAsp 562  
 DB 44 CACGGCAAGCCCGTTCTTCCGAGGAGACCCAGGAGAAGGAT 3

RESULT 10  
 AAT69342  
 ID AAT69342 standard; cDNA; 473 BP.

XX  
 AC AAT69342;  
 XX  
 DT 26-FEB-1998 (first entry)  
 XX  
 DE Murine metastatic nucleic acid sequence.  
 XX  
 KW Mouse; murine; tumour; cancer; metastatic sequence; detection;  
 KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;  
 KW hypertrophy; screening; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN W09718454-A2.  
 XX  
 PD 22-MAY-1997.  
 XX  
 PF 15-NOV-1996; 96WO-US18567.  
 XX

PR 30-JAN-1996; 96US-0594031.  
 PR 16-NOV-1995; 95US-0006838.  
 XX  
 PA (THOM/) THOMPSON T.  
 XX  
 PI Thompson T;  
 XX  
 XX WPI; 1997-289397/26.  
 XX  
 XX Identifying tumour metastatic sequences - by introducing transfected  
 PT cells into host mammal and analysing primary and metastatic  
 PT sequences by differential display PCR  
 XX  
 XX Disclosure; Fig 12DF; 102pp; English.  
 PS  
 XX Mouse Urogenital sinus (UGS) tissue was isolated from 17 day old  
 CC mouse embryos. The UGS cells were infected with retroviruses,  
 CC cultured and implanted under the renal capsule of mice.  
 CC Reconstructions were harvested 5 weeks later, when they showed  
 CC signs of distress from the tumour burden. Metastasised tumours were  
 CC isolated from a site outside the renal capsule. RNA was isolated  
 CC from primary tumours and metastases, reverse transcribed and  
 CC subjected to differential display PCR. The sequences were analysed  
 CC to obtain metastatic sequences, e.g. the present sequence. The  
 CC method can be used to detect, diagnose and treat disorders related  
 CC to metastasis, or treat malignant or non-malignant disorders, e.g.  
 CC hyperplasia, dysplasia and hypertrophy. The metastatic sequence can  
 CC be used to screen a biological sample for metastasis, and it or its  
 CC expression product may also be used to treat a metastatic disorder.  
 XX  
 SQ Sequence 473 BP; 117 A; 165 C; 85 G; 106 T; 0 other;

Alignment Scores:  
 Pred. No.: 4.54e-46 Length: 473  
 Score: 717.00 Matches: 136  
 Percent Similarity: 90.45% Conservative: 6  
 Best Local Similarity: 86.62% Mismatches: 13  
 Query Match: 23.42% Indels: 2  
 DB: 18 Gaps: 1

US-09-943-075A-2 (1-572) x AAT69342 (1-473)

QY 274 LeuAsnTyrSerAlaIleSerTyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheVal 293  
 DB 2 CTCACGACTCTGCCATTTCTTACAAAGTGGAACCTTGGGACACACTGGCTGTGTC 61  
 QY 294 SerAsnAsnHisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeu 313  
 DB 62 TCCACAATCACACTTTGAATCACACTTATGTGCTCAATGGAACCTTCAACCTTAACCTC 121  
 QY 314 ThrValGlnThrAlaValProGlyProCysProSerProThrProSer-----ProSer 331  
 DB 122 ACCGTCAAACTGCAGTGCCCGGGCATGCCCTCCCTTCGCTCGACTCGGCTCCA 181  
 QY 332 SerSerThrSerProSerProAlaSerSerProSerProThrLeuSerThrProSerPro 351  
 DB 182 CCTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTGCCCACATTTATCAACACCTAGCCCC 241  
 QY 352 SerLeuMetProThrGlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnCys 371  
 DB 242 TCTTTAATGCCTACTGGTTACAAATCCATGGAGCTGAGTGACATTTCCATGAAACCTGC 301  
 QY 372 ArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGlu 391  
 DB 302 CGAATAACAGATATGGCTACTTTCAGAGCCACCATCAATGATTGAGGGGATCTCGAA 361  
 QY 392 ValAsnIleIleGlnValAlaAspValProIleProThrLeuProAspAsnSerLeu 411  
 DB 362 GTGAGCATCATGCAGATAGCATGTCCCATGCGCCACCGAGCTGCCACTCCCTG 421  
 QY 412 MetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCysThr 428  
 DB 422 ATGGACTTCACTGTGACCTGCAAGGGGGCCACCCCATGGAAGCTGTACG 472

```

RESULT 11
AAQ96055
XX AAQ96055 standard; cDNA to mRNA; 2115 BP.
AC AAQ96055;
XX
XX 22-JAN-1996 (first entry)
XX
XX Sequence encoding melanoma associated antigen gp100.
XX
XX Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
XX Identification; tumour; gp100; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 22..2007
XX FT /*tag= a
XX FT /product= Melanoma associated antigen gp100.
XX FT 1..81
XX FT /*tag= b
XX FT 1792..1870
XX FT /*tag= c
XX FT /*label= Transmembrane domain.
XX FT 262..264
XX FT /*tag= d
XX FT /bound_moeity= Carbohydrate.
XX FT 337..339
XX FT /*tag= e
XX FT 352..354
XX FT /*tag= f
XX FT 982..984
XX FT /*tag= g
XX FT 1723..1725
XX FT /*tag= h
XX
XX EP668350-A1.
XX
XX 23-AUG-1995.
XX
XX 14-FEB-1995; 95EP-0200348.
XX
XX 21-DEC-1994; 94EP-0203709.
XX 16-FEB-1994; 94EP-0200337.
XX
XX (ALKU ) AKZO NOBEL NV.
XX
XX Adema GJ, Figdor CG;
XX
XX WPI: 1995-284790/38.
XX P-PSDB; AAR78646.
XX
XX Melanoma associated antigen gp100 - used in vaccines and for the
XX detection of tumours
XX
XX Claim 2; Page 19-22; 40pp; English.
XX
XX Immunogenic peptides derived from the melanoma associated antigen
XX may be used in the production of vaccines. Nucleotide sequences
XX encoding the immunogenic peptides may be used as primers and probes
XX in the detection of melanoma cells. Tumour infiltrating lymphocytes
XX capable of binding to the melanoma associated antigen can be
XX cultured ex vivo and returned to melanoma particles, and when
XX radiolabelled, they may be used to identify tumour deposits.
XX
XX Sequence 2115 BP; 469 A; 587 C; 575 G; 484 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 2,16e-38 Length: 2115
XX Score: 627.50 Matches: 176
XX Percent Similarity: 41.8% Conservative: 96
XX Best Local Similarity: 27.0% Mismatches: 205

```

```

Query Match: 20..50% Indels: 173
DB: 16 Gaps: 19
US-09-943-075a-2 (1-572) x AAQ96055 (1-2115)
QY 7 ValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGlnAlaAlaLys-----Arg 24
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 46 CTTTTCATTGCTGTGATAGGTGCTTGTGCTGGCTGGGCTACAAAAGTACCCAGA 105
QY 25 PheArgaspValLeuGlyHisGluGlnTyrProAspHisMetArgGluAsnGlnLeu 44
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 106 AACCAAGCAGCTGCTGTGT-----GTCTCAAGGCAACTC 138
QY 45 ArgGlyTyrSerSerAspGluAsnGluTyrAspGluGlnLeuTyrProValTyrArg 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 139 AGA-----ACCAAGCCTGGAACAGCAGCTGTATCCAGAGTGGACA--- 180
QY 65 GlyGluGlyArgTyrLysAspSerTyrGluGlyValGlnAlaAlaLeuThrSer 84
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 --GAAGCCCAAGACTTGACTGCTGGAGAGTGGTCAAGTGTCCCTCAAGTCAAGTAAT 237
QY 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeuValPheProArg 104
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 238 GATGGGCTACACTGATGTGCTGCAAAATGCTCTCTCTATTGCTTGAACCTCCCTGGA 297
QY 105 CysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsn---CysArgSerAsp 123
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 298 AGCCAAAAGGTATTGCCAGATGGGAGGTATTCTGGGTCAACAATACCATCAATCAATGGG 357
QY 124 LeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrGlyAlaAspAspGluAsp 143
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 358 AGCCAGGTGTGGGAGGAGCAGCAGTGTATCCCAAGAACT-----GACGATGCC--- 408
QY 144 TrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPhePro 163
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 409 -----TGCATCTTCCCTGATGGTGGACCTTGCCTG 438
QY 164 ArgProHisGlyArgLysLysTyrPsnPheValTyrValPheHisThrLeuGlyGlnTyr 183
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 439 TCTGGCTCTTGGTCTCAGAAAGAGAGCTTTGTTTATGTCTGGAAGACCTGGGGCCAATAC 498
QY 184 PheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsnLeuThrVal 203
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 499 TGGCAAGTTCTAGGGGGCCCGCAGTGTCTGGCTGAGCATTTGGGACAGCGGCAATGCTG 558
QY 204 GlyProGlnValMetGluValIleValPheArgArgHisGly---ArgAlaTyrIlePro 222
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 559 GGCACACACACCATGGAAGTGACTGTCTACCATCGCGGGGATCCCGAGCTATGCGCT 618
QY 223 IleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheValThrMet 242
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 619 CTGTTCATTCCAGCTCAGCCTTACCATTTACTGACCAAGTGGCTTCTTCCGTGACGCTG 678
QY 243 TyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhe 262
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 679 TCCCAAGTTGGCGGCTTGGATGGAGGCAACAACACTTCTCTGAGAAATCAGCTCTGACC 738
QY 263 PheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLys 282
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 739 TTTGGCTCCAGCTCCATGATGCCCCAGTGGCTATCTGGCTGAAGCTGACCTCTCTACACC 798
QY 283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 799 TGGGNCCTTGGAGACAGTAGTGAACCTGATCTCTCGGGGCACACTGTGTGCTCACTACT 858
QY 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro----- 320
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 859 TACCTGGAGCCTGCCCAAGTCACTGCTCCAGGTGCTCTGACGCTGCCATCTCTCTCACC 918
QY 321 -----GlyProCysPro----- 324
DB 919 TCCTGTGGCTCTCTCCCAAGTTCACAGGACCAACAGATGGGCACAGGCCAACTCCAGAGGCC 978

```

```
QY 324 ----- 324
Db 979 CCTAACACACAGCTGGCCAAAGTGCCTACTACAGAAAGTTGTGGTACTACACCTGGTCAG 1038
QY 325 SerProThrProSerSerSerTh----- 334
Db 1039 GCGCCNACTGCAGAGCCCTCTGGAAACACATCTGTGCAGGTGCCAACCACTGAAGTCATA 1098
QY 335 SerProSerProAlaserSerProSerProThrLeuSerThrProSerProSerLeuMet 354
Db 1099 AGCAGTGCACCTGTGCAGATGCCAACTGCAGACAGACAGGTATGACACCTGAGAAGGTG 1158
QY 355 Pro-----ThrGly 357
Db 1159 CAGGTTTCAGAGTCATGAGGTACACACTGGCAGAGATGTCAACTCCAGAGGCTCAGAGT 1218
QY 358 TyrLysSerMetGluLeuSerAspIle----- 366
Db 1219 ATGACACCTGCAGAGGTATCAATTTGTGGTCTTTCTGGAACACACAGCTGCACAGGTAA 1278
QY 366 ----- 366
Db 1279 ACTACAGAGTGGTGGAGACACACAGCTAGAGAGCTACCTATCCCTGAGCCTGAAGTCCA 1338
QY 367 -----SerAsnGlu----- 369
Db 1339 GATGCCAGCTCAATCATCTCTACGCAAAATATTACAGGTTCCTGGGCCCTGCTGGAT 1398
QY 370 -----AsnCysArgIleAsnArg 375
Db 1399 GGTACAGCCACTTAAGCGTGGTGAAGACAGACAGTCCCTCGGATTTCTGTATCGA 1458
QY 376 TyrGlyTyrPheArgAlaThrIleThrIleValAspGlyLeuGluValAsnIleIle 395
Db 1459 TATGTTCTCTTTCCTCGTCCACCTGGACATTTCTCAGGTTATTAAGTCCGAGATCCTG 1518
QY 396 GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
Db 1519 CAGGCT-----GTGCGCTCGGTCGGAGGGAT-----GCATTGAGCTGACT 1560
QY 416 ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435
Db 1561 GTGCTCTGCCAAGCGGGCTGCCAAGGAAGCCTGCATGGAGATCTCATCGCAGGGTGC 1620
QY 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValAspLulCysLeuLeuSer 455
Db 1621 CAGCCCCCTGCCAGCGCTGTGCCACCTGTGCTACCCAGCCAGCCCTGCCAGCTGGTT 1680
QY 456 ValArgArgAlaPhe---AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAsp 474
Db 1681 CTGCACCATGACTCAAGGGTGGCTCGGGGACATACCTGCTCAATGTCTCTGCTGAT 1740
QY 475 AspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGly--- 493
Db 1741 ACCACAGCCTGGCAGTGGTGCAGCCAGCCAGCTTATATGCTGCTGTCAAGACAGCCCTT 1800
QY 494 SerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheVal 513
Db 1801 GGGCAGGTTCGCTGATCGTGGCATCTTG-----CTGGTGTGTGATGGCT 1845
QY 514 ThrMetValThrIleLeuLeuTyrLysLys 523
Db 1846 GTGGTCTTGCATCTCTGATATATAGCGCG 1875
RESULT 12
ID AAH43500
XX AAH43500 standard; cDNA; 2130 BP.
AC AAH43500;
XX
DT 13-DEC-2001 (first entry)
XX
DE Human melanoma antigen gp100 coding sequence.
```

```
XX
KW Major histocompatibility complex; MHC; human; melanoma antigen; gp100;
KW HLA-A2 binding domain; mutation; antigen presenting cell; vaccine;
KW Immune effector cell; cancer; antibody; ss.
XX Homo sapiens.
XX
FH Key
CDS 22..2007
FT /*tag= a
FT /product= "gp100"
XX
XX WO200170767-A2.
XX
XX 27-SEP-2001.
XX
XX 19-MAR-2001; 2001WO-US08919.
XX
XX 20-MAR-2000; 2000US-190750P.
XX
XX 12-DEC-2000; 2000US-255019P.
XX
XX (GENZ ) GENZYME CORP.
XX
XX Nicolette CA;
XX
XX WPI; 2001-611469/70.
XX
XX P-PSDB; AAB47500.
XX
XX Novel synthetic compounds useful for stimulating an immune response in
XX a subject and as components of anti-cancer vaccines, are designed to
XX enhance binding to major histocompatibility complex molecules -
XX
XX Disclosure; Page 60-63; 67pp; English.
XX
XX This sequence encodes human melanoma antigen gp100. Peptides of the
XX invention based on the sequence of residues 209-217 of human melanoma
XX antigen gp100, which represents the putative HLA-A2 binding domain,
XX are designed to enhance binding to major histocompatibility complex
XX (MHC) molecules and to enhance immunoregulatory properties relative to
XX their natural counterparts. The mutations in the claimed peptides
XX confer tighter binding to the MHC. These peptides are useful for
XX inducing an immune response in a subject, where they are delivered in
XX the context of an MHC molecule which presents the compound on the
XX surface of an antigen presenting cell. The peptide sequences are useful
XX as components of anti-cancer vaccines and to expand immune effector cells
XX that are specific for cancers characterized by expression of the
XX melanoma antigen gp100. They are useful for diagnosis and treatment of
XX diseases such as cancer, in particular against human melanoma and for
XX generating antibodies that specifically recognize and bind the compounds.
XX
XX Sequence 2130 BP; 484 A; 587 C; 575 G; 484 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,18e-38 Length: 2130
Score: 627.50 Matches: 176
Percent Similarity: 41.85% Conservative: 96
Best Local Similarity: 27.08% Mismatches: 205
Query Match: 20.50% Indels: 173
DB: 22 Gaps: 19
US-09-943-075A-2 (1-572) x AAH43500 (1-2130)
QY 7 ValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGlnAlaAlaLys-----Arg 24
Db 46 CTCTTTCATTTGGCTGTGATAGTGTCTTGTGGGCTGTGGTGTGCTCAAAAGTACCCAGA 105
QY 25 PheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGluAsnAsnGlnLeu 44
Db 106 ARCCAGGACTGGCTTGGT-----GTCTCAAGGCAACTC 138
QY 45 ArgGlyTyrPheSerSerAspGluAsnGlnTrpAspGluGlnLeuTyrProValTrpArgArg 64
Db 139 AGA-----ACCAAGCCTTGAACAGGAGCTGTATCCAGATGGACA--- 180
```





PF 21-MAY-2001; 2001WO-US16417.  
 XX  
 PR 31-MAY-2000; 2000US-208955P.  
 PR 09-FEB-2001; 2001US-267877P.  
 XX  
 XX (GENZ ) GENZYME CORP.  
 PA  
 XX  
 PI Nicolette CA;  
 XX  
 XX WPI; 2002-106301/14.  
 DR P-PSDB; AAU09695.  
 XX  
 XX Novel anti-melanoma compound or peptide useful for inducing immune  
 PT response in a subject, for treating melanoma, as components of  
 PT anti-cancer vaccines and to expand immune effector cells specific for  
 PT cancers  
 XX  
 PS Disclosure; Page 64-65; 69pp; English.  
 XX  
 CC The present invention relates to anti-melanoma compounds comprising a  
 CC peptide sequence based on human melanoma antigen gp100. Also described  
 CC are antibodies that recognise and bind to these compounds,  
 CC polynucleotides that encode these compounds, and immune effector cells  
 CC that have been raised in vitro or in vivo in the presence of an antigen  
 CC presenting cell (APC) that presents the compound. Such an APC may be  
 CC the major histocompatibility complex (MHC) molecule. The anti-melanoma  
 CC compounds are useful for inducing an immune response in a subject, by  
 CC delivering the compound to the subject in the context of an MHC molecule  
 CC which presents the compound on the surface of an APC. The anti-melanoma  
 CC compound is delivered as a polynucleotide that encodes it. The compounds  
 CC are useful to generate antibodies that specifically recognise and bind  
 CC to them, for the treatment of melanoma, as components of anti-cancer  
 CC vaccines, and to expand immune effector cells that are specific for  
 CC cancers characterised by expression of gp100 tumour antigen, melanoma.  
 CC The compounds are also useful in diagnostic methods for such diseases.  
 CC The present sequence encodes human melanoma antigen gp100.  
 XX  
 SQ Sequence 2130 BP; 484 A; 587 C; 575 G; 484 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,18e-38 Length: 2130  
 Score: 627.50 Matches: 176  
 Percent Similarity: 41.85% Conservative: 96  
 Best Local Similarity: 27.08% Mismatches: 205  
 Query Match: 20.50% Indels: 173  
 DB: 24 Gaps: 19  
 US-09-943-075A-2 (1-572) x AAS14396 (1-2130)  
 QY 7 ValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGlnAlaAlaLys-----Arg 24  
 DB 46 CTTCTTCATTGGCTGCTGATAGGTGCTTTGCTGCTGCTGGGGCTACAAAAGTACCAG 105  
 QY 25 PheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGluAsnGlnLeu 44  
 DB 106 AACAGGAGCTGGCTGGT-----GTCTCAAGGCAAGCTC 138  
 QY 45 ArgGlyTyrSerSerAspGluAsnGluTyrAspGluGlnLeuTyrProValTyrArgArg 64  
 DB 139 AGA-----ACCAAGCCCTGGACAGCAGCTGTATCCAGAGTGGACA--- 180  
 QY 65 GlyGluGlyArgTyrPlyAspSerTrpGluGlyArgValGlnAlaLeuThrSer 84  
 DB 181 ---GAAGCCACAGACTTGGACTGTGGAGAGGTGGTCAAGTGTCCCTCAAGGTTCAGTAAT 237  
 QY 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeuValPheProArg 104  
 DB 238 GATGGCCTACACTGATTGGTGCAATGCTCTCTCTATTCCTTGCCTTGAACCTCCCTGCA 297  
 QY 105 CysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsn---CysArgSerAsp 123  
 DB 298 AGCCAAAAGGTTATGCCAGATGGGAGGTTATCTGGGTCAACAATACCATCATCAATGGG 357

QY 124 LeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAspAspGluAsp 143  
 DB 358 AGCCAGGTGTGGGAGGACAGCCAGTGTATCCCGAGGAACCT-----GACGATGCC--- 408  
 QY 144 TrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPhePro 163  
 DB 409 -----TGCATCTTCCTGATGGTGGACCTTGCCT 438  
 QY 164 ArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeuGlyGlnTyr 183  
 DB 439 TCTGGCTCTTGGTCTCAGAAGAGAGCTTTCTTTATGCTCTGGAAGACCTGGGCAATAT 498  
 QY 184 PheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsnLeuThrVal 203  
 DB 499 TGGCAAGTTCTAGGGGCCAGCTGTCTGGCTGAGCATTTGGGACAGCAGCAGGCAATGCTG 558  
 QY 204 GlyProGlnValMetGluValIleValPheArgArgHisGly---ArgAlaTyrIlePro 222  
 DB 559 GGCACACACACCATGGAAGTGACTGTCTACCATGCCGGGATCCCGGAGCTATGTGCT 618  
 QY 223 IleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheValThrMet 242  
 DB 619 CTTCCTTCATTCAGCTCAGCTTCCACCATTTACTGACAGGTGCTTCTCTCGTGAGGGTG 678  
 QY 243 TyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhe 262  
 DB 679 TCCAGTTTCGGGGCTTGGATGGAGGAACAAGCACTTCTCTGAGAAATCAGCCTGACC 738  
 QY 263 PheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLys 282  
 DB 739 TTTGCCCTCCAGCTCCATGACCCAGTGGCTATCTGGCTGAAGCTGACCTCTCTTACACC 798  
 QY 283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302  
 DB 799 TGGCACTTTGGAGACAGTAGTGAACCTCTCTCTGGGCACCTTGTGGTCACTCATACT 858  
 QY 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro----- 320  
 DB 859 TACCTGGAGCTGGGCCAGTCACTGCCAGGTGGTCTCTGAGGCTGCCATTCCTCTCTCACC 918  
 QY 321 -----GlyProCysPro----- 324  
 DB 919 TCCTGTGGCTCTCTCCCGCTTCCAGGACACACAGATGGGACAGGCCAAGTGCAGAGGCC 978  
 QY 324 ----- 324  
 DB 979 CCTAACACACAGCTGGCCCAAGTGCCTACTACAGAGTTGTGGGTACTACACCTGGTTCAG 1038  
 QY 325 SerProThrProSerProSerSerSerThr----- 334  
 DB 1039 GCGCCAACTGCGAGAGCCCTCTGGAACACATCTGTGCAGGTGCGCAACCACTGAAGTCATA 1098  
 QY 335 SerProSerProAlaSerSerProSerProThrLeuSerThrProSerProSerLeuMet 354  
 DB 1099 AGCACTGCACCTGTGCAGATGCCAACTGCAGAGACACAGCATATGACACCTCAGAGAGGTG 1158  
 QY 355 Pro-----ThrGly 357  
 DB 1159 CCAGTTTCAGAGGTTCATGGGTACCACACTGGCAGAGATGTCAACTCCAGAGGCTACAGGT 1218  
 QY 358 TyrLysSerMetGluLeuSerAspIle----- 366  
 DB 1219 ATGACACCTGCGAGAGGTATCAATTGTGGTGTCTTCTGGAACCAAGCTGCACAGGTANCA 1278  
 QY 366 ----- 366  
 DB 1279 ACTACAGAGTGGTGGAGACACACAGCTAGAGAGTACCTATCCCTGAGCCTGAAGGTCCA 1338  
 QY 367 -----SerAsnGlu----- 369  
 DB 1339 GATGCCAGCTCAATCATGTCTACGGAAGATATTACAGTTTCCCTGGGCCCCCTGTGCTGGAT 1398  
 QY 370 -----AsnCysArgIleAsnArg 375

Db 1399 GGTACAGCACCTTAAGCGTGGTGAAGACAGCAAGTCCCGCTGGATTGGTCTGTATCGA 1458  
 QY 376 TyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle 395  
 Db 1459 TATGGTTCCCTTTCCGTCACCCCTGGACATGTGTCAGGGGTATTGAAGTGGCGAGATCCTG 1518  
 QY 396 GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415  
 Db 1519 CAGCGT-----GTGCCGTCCGGTGAGGGGAT-----GCATTTGAGCTGACT 1560  
 QY 416 ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435  
 Db 1561 GTGTCTTGCACCAAGCGGGCTGCCCAAGGAAGCTGTCATGGAGATCTCATCGCCAGGGTGC 1620  
 QY 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSer 455  
 Db 1621 CAGCCCCCTGCCAGCGGCTGTGCCAGCCTGTGCTACCCAGCCCGCTGCCAGCTGGTT 1680  
 QY 456 ValArgArgAlaPhe---AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAsp 474  
 Db 1691 CTGCACAGATACTGAAGGTGGCTCGGGGACATAGTGCCTCAATGTGTCTGTGCTGAT 1740  
 QY 475 AspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGly--- 493  
 Db 1741 ACCAACAGCCTGGCAGTGGTCACGACCCAGCTTATCATGCTGCTGTCAGAGCAGGCTTT 1800  
 QY 494 SerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheVal 513  
 Db 1801 GGGCAGGTCCGCTGATCGTGGGCATCTTG-----CTGGTGTGTATGGCT 1845  
 QY 514 ThrMetValThrIleLeuLeuTyrLysLys 523  
 Db 1846 GTGGTCTTCNTCTCTGATATATAGGCGC 1875

RESULT 14  
 AAT02716  
 ID AAT02716 standard; cDNA; 2172 BP.  
 AC  
 XX  
 XX  
 XX  
 XX  
 DE MART-1 melanoma antigen cDNA25.  
 KW cDNA25; MART-1; melanoma antigen recognised by T-cells;  
 KW gp100 antigen derivative; melanoma; metastatic melanoma;  
 KW tumour-associated antigen; immunogen; diagnosis; prognosis;  
 KW prophylaxis; therapy; vaccine; ds.  
 XX  
 OS Mammalian.  
 XX  
 FH Key Location/Qualifiers  
 CDS 38..2038  
 FT /\*tag= a  
 FT /note= "cDNA25 melanoma antigen"  
 XX  
 XX W09529193-A2.  
 XX  
 XX 02-NOV-1995.  
 XX  
 PF 21-APR-1995; 95W0-US05063.  
 XX  
 XX 05-APR-1995; 95US-0417174.  
 PR 22-APR-1994; 94US-0231565.  
 XX  
 XX (USSH ) US SEC DEPT HEALTH.  
 PA  
 XX Kawakami Y, Rosenberg SA;  
 PI WPI; 1995-382963/49.  
 XX P-PSDB; AAR84854.  
 DR  
 XX

PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 XX  
 PS Disclosure; Fig 4A-4B; 184pp; English.  
 XX  
 CC The nucleic acid encodes cDNA25, a melanoma antigen (MART-1)  
 CC which is recognized by T-lymphocytes. cDNA25 is a derivative of  
 CC the melanocyte-melanoma-specific antigen gp100 (see AAR84855).  
 CC Antigen cDNA25 is a source of immunogenic peptides (see AAR84199)  
 CC which are optionally modified (see AAR84200-R84211) and used in  
 CC medicaments, especially vaccines, for the treatment or prevention  
 CC (by immunization) of melanoma. Antibodies against cDNA25 and its  
 CC immunogenic peptides may be used in the detection and isolation  
 CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).  
 XX  
 XX Sequence 2172 BP; 512 A; 594 C; 578 G; 488 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 2,44e-38 Length: 2172  
 Score: 627.00 Matches: 174  
 Percent Similarity: 41.03% Conservative: 96  
 Best Local Similarity: 26.44% Mismatches: 206  
 Query Match: 20.48% Indels: 182  
 DB: 16 Gaps: 19

US-09-943-075A-2 (1-572) x AAT02716 (1-2172)

QY 17 LeuProLeuGlnAlaAlaLysArgPheArgAspVal----- 28  
 Db 12 TTACCAATCGCGACCGGGAAGACACAAATGATCTGGTCTAAAGATGCCTTCTTCAT 71  
 QY 29 -----LeuGlyHisGluGlnTyrPro-----Asp 36  
 Db 72 TTGGCTGTGATAGGTGCTTTGCTGGCTGGGGCTACAAAGATACCCAGAACCCAGGAC 131  
 QY 37 HisMetArgGluAsnAsnGlnLeuArgGlyTyrPheSerSerAspGluAsnGluTyrPheAspGlu 56  
 Db 132 TGGCTTTGGTGTCTCAAGGCAACTCAGA-----ACCAAGCCTGGACAGG 176  
 QY 57 GlnLeuTyrProValTyrArgGlyGlyGlyArgTyrLysAspSerTyrPheGlyGly 76  
 Db 177 CAGCTGTATCCAGAGTGGACA-----GAAGCCAGAGACTTCACTGCTGGAGAGGTGT 230  
 QY 77 ArgValGlnAlaAlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPhe 96  
 Db 231 CAAGTGTCCCTCAAGGTCAAGTATGATGGGCTTACACTGATTGTCGCAATGCCTCCTTC 290  
 QY 97 ValValAsnLeuValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyr 116  
 Db 291 TCTATTGCTTGAACCTTCCCTGGGAAGCCAAAGGTATTGCCAGATGGGAGGTATTCTGG 350  
 QY 117 GluArgAsn---CysArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTyr 135  
 Db 351 GTCAACAATACCATCATCAATGGGAGCCAGGTGTGGGGAGGACAGCCAGGTATCCCCAG 410  
 QY 136 ThrThrGlyAlaAspAspGluAspTyrPheGluAsnThrSerGlnGlyGlnHisLeuArg 155  
 Db 411 GAAACT-----GACGATGCC-----TGCATC 431  
 QY 156 PheProAspGlyLysProPheProArgProHisGlyArgLysLysTyrPheAsnIleValTyr 175  
 Db 432 TTCCCTGATGTGGACCTTGCCCATCTGGCTCTTGGTCTCAGAAGAGAAAGCTTTGTTTAT 491  
 QY 176 ValPheHisThrLeuGlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSer 195  
 Db 492 GTCTGGAAGACCTGGGGCCAAATCTGGAATTTCTAGGGGGGCCAGTGTCTGGGCTGAGC 551  
 QY 196 IleAsnThrValAsnLeuThrValGlyProGlnValMetGluValIleValPheArgArg 215  
 Db 552 ATTGGGACAGGCGGCAATGCTGGGCACACACACACCATGGAAGTACTGCTACCATCCG 611

```
QY 216 HisGly---ArgAlaTyrIleProIleSerLysValIleThrAsp 234
    ||| |||:||||:||||:||||: ||| ||| ||| ||| ||| ||| |||
Db 612 CGGGATCCCGGAGCTATGTGCTCTTCTATTCCAGCTCAGCCTTTCACATTACTGAC 671
QY 235 GlnIleProIlePheValThrMetTyrGlnLysAsnAspArgAsnSerSerGluThr 254
    |||:|||| |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
Db 672 CAGTGCCCTTCTCCGTGAGCTGTCCAGTTCGGGCTTGGATGGAGGAACAGCAC 731
QY 255 PheLeuArgAspLeuProIlePhePheAspValLeuIleHisAspProSerHisPheLeu 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 732 TTCTTCAGAAATCAGCCTCTGACCTTTGCTCCAGCTCCATGACCCAGCTATCTG 791
QY 275 AsnTyrSerAlaIleSerTyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSer 294
    |||:|||| |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
Db 792 CTTGAAGCTGACCTCTCTACACCTGGGACTTTGGACACAGTAGTGAACCCCTGATCTCT 851
QY 295 AsnAsnHisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThr 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 852 CGGGCACTTTGTGTCACTACTACTTACCTGGAGCCTGGCCAGCTCACTGGCCAGGTGTC 911
QY 315 ValGlnThrAlaValPro-----GlyProCysPro----- 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 912 CTGCAAGCTGCCATTCCTCTACCTCTGCTGCTCTCCCTCCAGTTCAGGACACAGAT 971
QY 324 ----- 324
Db 972 GGGCACAGGCCAACTGCAGAGGCCCTAACACACAGCTGGCCAAAGTGCCTACTACAGAA 1031
QY 325 -----SerProThrProSerProSerSerSerThr----- 334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1032 GTTGTGGGTACTACACCTGGTCAGCGCCAACTGCAGAGCCCTCTGGAACACATCTGTG 1091
QY 335 -----SerProSerProAlaSerSerProSerProThrLeu 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1092 CAGGTGCCAACCACTGAAGTCTAAGACACTGCACCTGTGCAGATGTCACCACTGCAGAGC 1151
QY 347 SerThrProSerSerLeuMetPro----- 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1152 ACAGGTATCACACCTCAGAGAGTGCAGAGTTCATGGGTACACACACTGGCAGAG 1211
QY 356 -----ThrGlyTyrLysSerMetGluLeuSerAspIle----- 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1212 ATGTCAACTCCAGAGGCTACAGGTATGACACCTGCAGAGGTATCAATTTGGTCTTCT 1271
QY 366 ----- 366
Db 1272 GGAACCAAGCTGCACAGGTAACTACAGAGTGGTGGAGACCACTAGAGAGCTA 1331
QY 367 -----SerAsnGlu----- 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1332 CCTATCCCTGAGCTGAAGTCCAGATGCCAGCTCAATCATGTCTACGGAAGTATTACA 1391
QY 369 ----- 369
Db 1392 GGTTCCTCGGCCCTCGTGGATGGTACACCCACTTAAGCTGGTGAAGACAGCAAGTC 1451
QY 370 -----AsnCysArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAsp 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1452 CCCTCGATTGCTTCTGTATGATATGTTCTTTCCTCCGTCACCCCTGGACATTTGTCAG 1511
QY 388 GlyIleLeuGluValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnPro 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1512 GCTATTGAAGTCCCGAGATCTCTGCAAGCT-----GTGCGCTCCGCTGAGGGG 1559
QY 408 AspAsnSerLeuMetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCys 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1560 GAT-----GCATTTGAGCTGACTGTCTCTGCCAAGCGGGTGCACCAAGAGCCTGC 1613
QY 428 ThrIleIleSerAspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAla 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1614 ATGGAGATCTCATCGCGAGGTGCCAGGCCCTGCTGCCAGCGCTGTGCTA 1673
QY 448 ValAspGluLeuCysLeuLeuSerValArgAlaPhe---AsnGlySerGlyThrTyr 466
```

```
Db 1674 CCGAGCCCGCCGCGCCAGCTGGTCTTGCCACAGATACCTGAAGGGTGGCTCGGGGACATAC 1733
QY 467 CysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuIleSer 486
    |||:|||| |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
Db 1734 TGCTCAATGTGCTCTCGCTGATACCAACAGCCCTGGCAGTGGTCAAGCACCAGCTTATC 1793
QY 487 IleProGlyLysAspLeuGly---SerProLeuArgThrValAsnGlyValLeuIleSer 505
    |||:|||| |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
Db 1794 ATGCTGGTCAAGAAGCAGCGCTTGGCAGGTTCGCTGATCGTGGCGCATCTTG----- 1847
QY 506 IleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeuTyrLysLys 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1848 -----CTGCTGTGTGATGCTGTGGTCCCTTGCATCTCTGATATATAGCGC 1892

RESULT 15
AAS45525
ID AAS45525 standard; cDNA; 2172 BP.
XX
AC AAS45525;
XX
XX 18-DEC-2001 (first entry)
XX
DE DNA encoding Melanoma antigen cDNA25.
XX
KW Human; MART-1; immunogenic; melanoma antigen recognised by T lymphocyte;
KW diagnostic; therapeutic; vaccine; melanoma; in vivo tumour recognition;
KW in vivo tumour rejection; ss.
XX
OS Homo sapiens.
XX
XX US6270778-B1.
XX
XX 07-AUG-2001.
XX
XX 12-MAR-1999; 99US-0267439.
XX
XX 05-MAY-1998; 98US-0073138.
XX
XX 22-APR-1994; 94US-0231565.
XX
XX 03-APR-1995; 95US-0411714.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Kawakami Y, Rosenberg SA;
XX
XX WPI; 2001-595403/67.
XX
XX P-PSDB; AAU28912.
XX
XX Immunogenic peptide useful in vaccines comprises specific amino acids
XX of new melanoma antigen recognised by T lymphocytes.
XX
XX Example 3; Figure 4; 73pp; English.
XX
XX The invention relates to a novel immunogenic peptide comprising 5-20
XX contiguous amino acids of new melanoma antigen recognised by T
XX lymphocytes (MART-1). The peptide sequence contains at least one amino
XX acid modification of MART-1. The peptide is used in diagnostic and
XX therapeutic methods as an immunogen or vaccine to prevent or treat
XX melanoma, and for in vivo tumour recognition and rejection. AAS4524-
XX AAS4528 represent MART-1 coding sequences, PCR primers, and related
XX sequences of the invention.
XX
XX Sequence 2172 BP; 512 A; 594 C; 578 G; 488 T; 0 other;

Alignment Scores:
Pred. No.: 2,44e-38 Length: 2172
Score: 627.00 Matches: 174
Percent Similarity: 41.03% Conservatve: 96
Best Local Similarity: 26.44% Mismatches: 206
Query Match: 22 Indels: 182
DB: Gaps: 19

US-09-943-075A-2 (1-572) x AAS45525 (1-2172)
```

[illegible]

Search completed: April 8, 2003, 14:52:59  
Job time : 381 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 9, 2003, 02:18:55 ; Search time 69 seconds  
(without alignments)  
2542.306 Million cell updates/sec

Title: US-09-943-075A-2

Perfect score: 3061

Sequence: 1 MESLCGVLVFLAAGLPLQ.....PFRSGDREKPLLDKPMWL 572

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09943075/runat\_07042003\_155631\_17949/app\_query.fasta\_1.711  
-DB=issued\_patents\_NA -OFF=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRB=US09943075.ecgn\_1\_1\_40.runat\_07042003\_155631\_17949 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAB -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2700	88.2	2213	4 US-09-383-586-27	Sequence 27, Appl
2	2163	70.7	2669	1 US-08-594-031-91	Sequence 91, Appl
3	2163	70.7	2669	1 US-08-594-031-101	Sequence 101, Appl
4	941	30.7	620	1 US-08-594-031-113	Sequence 113, Appl
5	729	23.8	524	1 US-08-594-031-105	Sequence 105, Appl
6	717	23.4	473	1 US-08-594-031-115	Sequence 115, Appl
7	627	5	20.5	4 US-09-056-105-1	Sequence 1, Appl
8	627	20.5	2172	2 US-08-417-174-26	Sequence 26, Appl
9	627	20.5	2172	2 US-08-231-565A-26	Sequence 26, Appl
10	627	20.5	2172	2 US-08-007-961-26	Sequence 26, Appl
11	627	20.5	2172	4 US-09-267-439-26	Sequence 26, Appl
12	524	17.1	335	1 US-08-594-031-92	Sequence 92, Appl

13	524	17.1	335	1	US-08-594-031-94	Sequence 94, Appl
14	524	17.1	335	1	US-08-594-031-96	Sequence 96, Appl
15	524	17.1	335	1	US-08-594-031-123	Sequence 123, Appl
c 16	504	16.5	354	1	US-08-594-031-114	Sequence 114, Appl
c 17	471	15.4	286	1	US-08-594-031-121	Sequence 121, Appl
18	477	13.9	262	1	US-08-594-031-93	Sequence 93, Appl
19	418	13.7	309	1	US-08-594-031-106	Sequence 106, Appl
c 20	479	13.4	280	1	US-08-594-031-119	Sequence 119, Appl
c 21	433	13.2	292	1	US-08-594-031-107	Sequence 107, Appl
c 22	388	12.7	270	1	US-08-594-031-109	Sequence 109, Appl
c 23	364.5	11.9	335	1	US-08-594-031-111	Sequence 111, Appl
c 24	356	11.6	223	1	US-08-594-031-116	Sequence 116, Appl
c 25	350	11.4	240	1	US-08-594-031-112	Sequence 112, Appl
c 26	317.5	10.4	217	1	US-08-594-031-112	Sequence 112, Appl
c 27	294	9.6	190	1	US-08-594-031-95	Sequence 95, Appl
c 28	293	9.6	231	1	US-08-594-031-120	Sequence 120, Appl
c 29	279	9.1	247	1	US-08-594-031-117	Sequence 117, Appl
c 30	216	7.1	224	1	US-08-594-031-122	Sequence 122, Appl
c 31	135	4.4	18596	4	US-09-318-448-11	Sequence 11, Appl
c 32	131.5	4.3	3004	1	US-08-276-213-6	Sequence 6, Appl
c 33	130	4.2	2293	1	US-08-604-913B-12	Sequence 12, Appl
c 34	129	4.2	13807	4	US-09-052-469-7	Sequence 5, Appl
c 35	129	4.2	14148	4	US-09-052-469-7	Sequence 7, Appl
c 36	128.5	4.2	377	2	US-08-332-766A-1	Sequence 1, Appl
c 37	126.5	4.1	3337	1	US-08-072-610-1	Sequence 1, Appl
c 38	126.5	4.1	3337	2	US-08-719-822B-1	Sequence 1, Appl
c 39	126.5	4.1	3337	4	US-09-092-458-1	Sequence 1, Appl
c 40	126	4.1	2900	1	US-08-034-650-9	Sequence 9, Appl
c 41	126	4.1	2900	1	US-08-449-015-9	Sequence 9, Appl
c 42	125	4.1	14060	3	US-08-658-136-4	Sequence 4, Appl
c 43	124	4.1	12912	2	US-08-460-751-1	Sequence 1, Appl
c 44	123	4.0	11707	4	US-09-136-574A-1	Sequence 1, Appl
c 45	120	3.9	4800	3	US-09-106-638-1	Sequence 1, Appl

# ALIGNMENTS

RESULT 1  
US-09-383-586-27  
; Sequence 27, Application US/09383586  
; Patent No. 6242419  
; GENERAL INFORMATION:  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Abernethy, Nevin  
; APPLICANT: Onrust, Rene  
; APPLICANT: Kumble, Anand  
; APPLICANT: Marison, Greg  
; TITLE OF INVENTION: Compounds isolated from stromal cells  
; TITLE OF INVENTION: and methods for their use  
; FILE REFERENCE: 11000.1037cl  
; CURRENT APPLICATION NUMBER: US/09/383.586  
; CURRENT FILING DATE: 1999-08-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 27  
; LENGTH: 2213  
; TYPE: DNA  
; ORGANISM: Mouse  
US-09-383-586-27

Alignment Scores:  
Pred. No.: 7.91e-257 Length: 2213  
Score: 2700.00 Matches: 507  
Percent Similarity: 92.29% Conservat: 20  
Best Local Similarity: 88.79% Mismatches: 42  
Query Match: 88.21% Indels: 2  
DB: 4 Gaps: 1

US-09-943-075A-2 (1-572) x US-09-383-586-27 (1-2213)

QY 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGln 20  
|||||

Db 23 ATGGAAGTCTCTGGGGTCTGGGATTTCTGCTGCTGGCTGGCTGAGGACTGCCTCTCCAG 82  
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluInTyrProAspHisMetArgGlu 40  
Db 83 GCTGCCAAGGCGATTTCTGATGCTGTGGGCCATGAACAGTATCCCAATCATCATGAGAG 142  
Qy 41 AsnAsnGlnLeuArgGlyTyrPheSerAspGluAsnGluTyrAspGluGlnLeuTyrPro 60  
Db 143 CACAACCAATTACGTGGCTGGCTTCGGATGAAATGAATGGATGAACACCTGTATCCA 202  
Qy 61 ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla 80  
Db 203 GTGTGGAGGAGGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 262  
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
Db 263 GTCCTGACCACTGACTCACCGCTCTGGTGGGTTCCTAATATCACCTTTGGTGAACCTG 322  
Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
Db 323 GTGTTCCTCCAGATGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 382  
Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAsp 140  
Db 383 AGGAATGATTGGGACTGACTCTGACCTGCTGCTACAACTGGACTGGAGGGGAGAT 442  
Qy 141 AspGluAspTrpGluAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
Db 443 GATGGTACCTGGGAAGATGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502  
Qy 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
Db 503 CCTTTCCTCGCCCGATGGATGGAGAAATGGAGCTTTGCTACGCTTTTCACACACTT 562  
Qy 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
Db 563 GGCCAGTATTTCAAAATACTGGTGGTGTTCAGCACCGGTTTCTATAAACAAGTCAAC 622  
Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220  
Db 623 TTGACAGCTGGCCCTCAGGTCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 682  
Qy 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
Db 683 ATTCCTCATCTCGAAGTGAAGATGTATGTATGATAACAGATCAGATCCCTGTATCGTG 742  
Qy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
Db 743 ACCATGTCACAGAAGATGACAGGAACCTGTCTGATGAGATCTTCTCAGAGACCTCCCC 802  
Qy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
Db 803 ATCGTCTTCGATGTCCTATTCATGATCCAGCCACTTCCCTCAACGACTCTGCCATTTC 862  
Qy 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn 300  
Db 863 TCAAGTGGAACTTGGGACAACTGGCCCTGTGTCTCCAAATCACACTTGAAT 922  
Qy 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
Db 923 CACACTTATGTGCTCAATGAACCTTCAACCTTAACCTCAACCTGCAAACTGCAAGTCC 982  
Qy 321 GlyProCysProSerProThrProSer-----ProSerSerSerThrSerProSerPro 338  
Db 983 GGGCCATGCCCTCCCTCTGGCTTCGACTCCGCTCCACCTTCAACTCCGCTCCACTT 1042  
Qy 339 AlaSerSerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 358  
Db 1043 CCGCCCTCACCTCTGCCCATATATCAACACTAGCCCTCTTTAATGCCCTACTGCTAC 1102  
Qy 359 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 378  
Db 1103 AAATCCATGAGCTGAGTGACATTTCCATGAAACTGCCCAATAAACAAGATATGGCTAC 1162

Qy 379 PheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAla 398  
Db 1163 TTTCAGAGCCCACTCAATATTAGAGGGATCTCTGGAAGTCAGCATCATCATGAGATGCA 1222  
Qy 399 AspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCys 418  
Db 1223 GATGTCCCATGCCACACCGAGCTGCCACTCCCTGATGGACTTCCTGCTGACCTGC 1282  
Qy 419 LysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 438  
Db 1283 AAAGGGGCCACCCCACTGGAAGCTGTAGATCATCTCCGACCCACCTGCCAGATGCC 1342  
Qy 439 GluAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArg 458  
Db 1343 CAGAACCGGCTGTGCAGCCCTGTGCTGTGGATGGCTGTGCTGTCTGTGAGAAGA 1402  
Qy 459 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 478  
Db 1403 GCCTTCAATGGGTCTGGCACCTACTGTGTGAATTTCATCTGGGAGATGATGCAAGCTG 1462  
Qy 479 AlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThr 498  
Db 1463 GCCTCACACGACCCCTGATCTATCCCTGGCAAAAGACCCAGACTCCCTCTGAGAGCA 1522  
Qy 499 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIle 518  
Db 1523 GTGAATGGTCTCTGATCTCCATTTGGCTGGCTGGCTGTGCTGTACCATGGTTACCATC 1582  
Qy 519 LeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValAl 538  
Db 1583 TTGCTGTACAAAAACACAAAGCGGTACAAAGCAATAGAAAACCTGCCCAAGAACACG 1642  
Qy 539 LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAsp 558  
Db 1643 AAGGCAAAAGGCTGTGAGTCTCTCTCAGCCACGCGAAAGCCCTTCTCCGAGGAGAC 1702  
Qy 559 ArgGluLysAspProLeuLeuGlnAspLysPro 569  
Db 1703 CAGGAGAAGGATCATTTGCTCCAGGACAAGCCA 1735

## RESULT 2

US-08-594-031-91  
: Sequence 91, Application US/08594031  
: Patent No. 5783182  
: GENERAL INFORMATION:  
: APPLICANT: THOMPSON, Timothy C.  
: TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
: NUMBER OF SEQUENCES: 175  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: BAKER & BOTTS, L.L.P.  
: STREET: 1299 Pennsylvania Avenue, N.W.  
: City: Washington  
: STATE: DC  
: COUNTRY: USA  
: ZIP: 20004-2400  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FastSeq Version 1.5  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/594,031  
: FILING DATE: 30-JAN-1996  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 60/006,838  
: FILING DATE: 16-NOV-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Remenick, James  
: REGISTRATION NUMBER: 36,902  
: REFERENCE/DOCKET NUMBER: 0A146-0110  
: TELECOMMUNICATION INFORMATION:



TELEPHONE: 202-639-7700  
TELEFAX: 202-639-7890  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2669 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-594-031-91

Alignment Scores:  
Pred. No.: 1, 24e-203 Length: 2669  
Score: 2163.00 Matches: 394  
Percent Similarity: 82.57% Conservative: 75  
Best Local Similarity: 69.37% Mismatches: 85  
Query Match: 70.66% Indels: 14  
DB: 1 Gaps: 1

US-09-943-075A-2 (1-572) x US-08-594-031-91 (1-2669)

QY 1 MetGluSerLeuGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln 20  
DB 92 ATGGAAATGCTCTACTATTTCCTGGGATTTCTGCTCGCTGGCTGCAAGATTGCCACTTGAT 151  
QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
DB 152 GCCGCCAAACGATTCATGATGCTGCTGGCAATGAAAGACCTTCTGCTACATGAGGAG 211  
QY 41 AsnAsnGlnLeuArgGlyTyrPheSerAspGluAsnGlnTyrPheValPheValAsnLeu 60  
DB 212 CACAATCAATTAATGGCTGGTCTCTGATGAAATGCAATGCAATGCAATGCAATGCAATGCAAT 271  
QY 61 ValTyrArgGlyGluGlyArgTyrPheLysAspSerTrpGluGlyArgValGlnAla 80  
DB 272 GTGTGGAGCGGGAGAGATGAGTGGTGGAAACCTCTGGAAGGAGGCGGTGTCAGGCG 331  
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
DB 332 GTCTGACAGTGAATCACCAGCCCTGCTGGCTCAATATATACATTTTGGCGTGAACCTG 391  
QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
DB 392 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 451  
QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrThrGlyAlaAsp 140  
DB 452 AGAATAGGCTGGTTTATCTGCTGATCCATATGTTTACACTGGACAGCATGGTGCAGAG 511  
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
DB 512 GACAGTGAAGGGAATGGCCGCAAGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 571  
QY 161 ProPheProArgProHisGlyArgGlyLysTyrAsnPheValTyrValPheHisThrLeu 180  
DB 572 CTTTTCTCACCACCCCGATGGAGATGGATGGAATTTCTATCTCTTCCACACACTT 631  
QY 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
DB 632 GGTCAAGTATTCAGAAATGGGACGATGTCAGTGAAGTTCCTGTGCAACACAGCCCAAT 691  
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220  
DB 692 GTGACACTTGGGCTCAACTCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 751  
QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
DB 752 GTTCCCATCGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 811

QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
DB 812 ACTATGTTCCAGAAAGACGATCGAAATTCATCGACGAAACCTTCTCTCAAGATCTCC 871  
QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
DB 872 ATTTATGTTTGGTCTGATTCATGATTCATGATTCATGATTCATGATTCATGATTCATGATTCAT 931  
QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
DB 932 TACAAGTGGAGCTTCGGGGATAATACTAGTGGCTGTTTGTTCACCAATCATCATCTGTAAT 991  
QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
DB 992 CACAGTATGTGCTCAATGCAACCTTCAGCTTAACCTCCTGCTGCAAGCTGCAGCACCA 1051  
QY 321 GlyProCysProSerProThrProSerProSerProSerProSerProSerProSerProSer 340  
DB 1052 GGAAGCTTGTGGCCAGCCGCCACCCACCCAGACCT----- 1087  
QY 341 SerProSerProThrLeuSerThrProSerProSerProSerProSerProSerProSerProSer 360  
DB 1088 -----TCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACCAACCCC 1129  
QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380  
DB 1130 CTGGAGCTGAGTAGGATTTCTGATGAAACTGCCAGATTAACAGATATGGGCCACTTTCAA 1189  
QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400  
DB 1190 GCCACCATCACAAATGTAGAGGAACTTTAGAGGTTAAACATCATCCAGATGACAGACGTC 1249  
QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420  
DB 1250 CTGATGCGGCGGCTGCAATGGCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGG 1309  
QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440  
DB 1310 AGCATTTCCACGGAGGCTGCTACCATCATTTCTGACCCACCTGCCAGATCACCCAGAAC 1369  
QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuSerValArgArgAlaPhe 460  
DB 1370 ACAATGCTCCAGCCCTGGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 1429  
QY 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeuAlaLeu 480  
DB 1430 AATGGTCTGGGAGCTGCTGTGAACCTCACCTGGGGATGACACAAAGCTGGCTCTC 1489  
QY 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500  
DB 1490 AGGAGCACCTGATTTCTGCTGACAGAGACCCAGCTTCCGCTTTAAGGATGCCAAC 1549  
QY 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
DB 1550 AGTGGCTGATCTCCGTTGGCTGCTGGCCATATTTGCTACCTGTGATCTCCCTCTGTGTG 1609  
QY 521 TyrLysLysHisThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540  
DB 1610 TACAAAAACACAGGAATACACCCCAATAGAAAATAGTCTGGGAAATGTGTGTGAGAAC 1669  
QY 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560  
DB 1670 AAAGGCTGTGCTGCTTCTCAACCGGTCAAAAGCGGTGTTCTTCCCGGAAACACGAGAA 1729  
QY 561 LysAspProLeuLeuGlnAspLys 568  
DB 1730 AAGGATCGCTACTCAAAACCAA 1753

RESULT 3  
US-08-594-031-101  
; Sequence 101, Application US/08594031  
; Patent No. 5783182

```

; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-594-031-101

Alignment Scores:
Pred No.: 1-24e-203 Length: 2669
Score: 2163.00 Matches: 294
Percent Similarity: 82.57% Conservative: 75
Best Local Similarity: 69.37% Mismatches: 85
Query Match: 70.66% Indels: 14
DB: 1 Gaps: 1

US-09-943-075A-2 (1-572) x US-08-594-031-101 (1-2669)
QY 1 MetGluSerLeuGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln 20
Db 92 ATGGAAATGCTCTACTATTTCCTGGGATTTCTGCTCGCTGCAAGATTGCCACTTGAT 151
QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
Db 152 GCCGCCAACGATTTTCATGATGCTGGGCAATGAAAGACCTTCCTGCTTACATGAGGGAG 211
QY 41 AsnAsnGlnLeuArgGlyTyrSerSerAspGluAsnGluTyrAspGluGlnLeuTyrPro 60
Db 212 CACATCAATTAATGCTGGTCTCTGATGAAATGACTGGAATGAAATACTCTACCCA 271
QY 61 ValTrpArgGlyGlyGluArgTyrLysAspSerTrpGluGlyGlyArgValGlnAla 80
Db 272 GTGTGGAAGCGGGAGACATGAGTGGCAAACTCTCTGGAAGGAGCGCGGTGTCAGGCG 331
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 332 GTCTTGACCACTGACTCACCAGCCCTCGTGGGCTCAAAATATAACATTTTGCCTGAACTG 391

```

```

QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAspCys 120
Db 392 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGATGCG 451
QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAsp 140
Db 452 AGAAATGAGGCTGTTTATCTGCTGATCATATGTTTACAACATGGACAGCATGGTCAGAG 511
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
Db 512 GACAGTGACGGGAAATGGCCAAAGCCATCAACGCTCTCCCTGATGGGAA 571
QY 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180
Db 572 CCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTAGCTCTTCACACACTT 631
QY 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Db 632 GGTCAAGTATTTCCAGAAATTTGGGACCATGTTCTAGTGAGAGTTTCTGTGAACACACCAAT 691
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220
Db 692 GTGACACTTGGCCCTCAACTCATGGAAGTACTGCTTACAGAAGACATGGAGCGCATAT 751
QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240
Db 752 GTTCCCATCGCACAAAGTGAAGATGTTAGTGTGTAAACAGATCAGATTCCTGTGTTCTG 811
QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
Db 812 ACTATGTTCCAGAAAGACGATCGAAATTCACGACGAAACCTCTCTCAAGATCTCCCG 871
QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
Db 872 ATTATGTTGATGCTCTGATTTCATGATCTCTAGCCACTTCTCAATTAATCTACCAATTAAC 931
QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
Db 932 TACAAGTGGAGCTTCGGGGATAATACACTGCGCTGTGTTGTTCCACCAATCATACTGTGAAT 991
QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
Db 992 CACACGTATGCTCAATGGAACCTTCAGCCTTAACTCCTGTAAGAGTGCAGCACCA 1051
QY 321 GlyProCysProSerProThrProSerProSerSerSerThrSerProSerProAlaSer 340
Db 1052 GCACCTGTCCGCCACCCGCCACCCACCCACCCAGACCT----- 1087
QY 341 SerProSerProThrLeuSerThrProSerProSerProSerProThrGlyTyrLysSer 360
Db 1088 -----TCAAAACCCACCCCTTCTTAGGACCTGCTGGTGCAACCC 1129
QY 361 MetGluLeuSerAspIleSerAsnGluAsnGlyArgIleAsnArgTyrGlyTyrPheArg 380
Db 1130 CTGAGCTGATGATGATTCCTGATGAAACTGCCAGATTAAACAGATATGGCCACTTTCAA 1189
QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleValAlaAlaAspVal 400
Db 1190 GCCACCATCAAAATGTAGAGGGAATCTTAGAGGTTAATCATCATCCAGATGACAGACGTC 1249
QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
Db 1250 CTGATGCCGCTGCCATCGCTGAAAGCTCCCTAATAGACTTGTGTGACCTGCCACAGG 1309
QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440
Db 1310 AGCATTCACCGGAGTCTGTACCATCATTTCTGACCCACCCACCTGGGAGATCACCAGAAC 1369
QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460
Db 1370 ACAGTCTGCACCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACCACTTC 1429

```

QY 461 AsnGlySerGlyThrTyrCysValAsnPhaThrLeuGlyAspAlaSerLeuAlaLeu 480  
 Db 1430 AATGGGTGGGACCTACTGTGAACCTCACCTGGGGATGACACAAAGCTGGCTCTC 1489  
 QY 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500  
 Db 1490 ACGAGCACCTGATTTCTGTCTTCAGAGAGACCCAGCCTCGCCTTTAAGGATGGCAAC 1549  
 QY 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
 Db 1550 AGTGCCCTGATCTCGTGTGGCTGCTGGCCATATTTGTCACTGTGATCTCCCTCTGTGGTG 1609  
 QY 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540  
 Db 1610 TACAAAAACACAGGAATACAAACCAATAGAAAATAGTCCCTGGGAATGTGTGTCAGAAAGC 1669  
 QY 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560  
 Db 1670 AAGGCGCTGAGTGTCTTCTCAACCGTGCAAAAGCGTGTCTTCCCGGAACACGAA 1729  
 QY 561 LysAspProLeuLeuGlnAspLys 568  
 Db 1730 AAGGCGCTGACTACTCAAAAACCAA 1753

RESULT 4

US-08-594-031-113  
 ; Sequence 113, Application US/08594031  
 ; Patent No. 5783182  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THOMPSON, Timothy C.  
 ; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
 ; NUMBER OF SEQUENCES: 175  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: BAKER & BOTTS, L.L.P.  
 ; STREET: 1299 Pennsylvania Avenue, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20004-2400  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq Version 1.5  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/594,031  
 ; FILING DATE: 30-JAN-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/006,838  
 ; FILING DATE: 16-NOV-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Remenick, James  
 ; REGISTRATION NUMBER: 36,902  
 ; REFERENCE/DOCKET NUMBER: 0A146-0110  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-639-7700  
 ; TELEFAX: 202-639-7890  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 113:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 620 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FRAGMENT TYPE:  
 ; ORIGINAL SOURCE:  
 ; US-08-594-031-113

Alignment Scores:

Pred. No.: 6,27e-84 Length: 620  
 Score: 941.00 Matches: 183  
 Percent Similarity: 91.75% Conservative: 6  
 Best Local Similarity: 88.83% Mismatches: 15  
 Query Match: 30.74% Indels: 3  
 DB: 1 Gaps: 1  
 US-09-943-075A-2 (1-572) x US-08-594-031-113 (1-620)  
 QY 274 LeuAsnTyrSerAlaIleSerTyrLysTrpAsnPhaThrGlyLeuPheVal 293  
 Db 2 CTCACAGACTCTGCCATTTCTCAAGTGGAACTTGGGACACACTGGCCTGTTTGTTC 61  
 QY 294 SerAsnAsnHisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPhaAsnLeu 313  
 Db 62 TCCAACAATCAACACTTTGAATCACACTTATGTGTCTCAATGGAACCTTCAACCTTAACTTC 121  
 QY 314 ThrValGlnThrAlaValProGlyProCysProSerProThrProSer-----ProSer 331  
 Db 122 ACCGTGCAAACTGCAAGTCCCGGGCCATGCCCTCCCTTCGCTTCGACTCGGCTCCCA 181  
 QY 332 SerSerThrSerProSerProAlaSerSerProSerProThrLeuSerThrProSerPro 351  
 Db 182 CCTTCAACTCGCGCCTCACCTCCGCTCCCTCACCTCTGCCCCACATTATCAACACCTAGCCCC 241  
 QY 352 SerLeuMetProThrGlyTyrLysSerMetGluLeuSerAspLysSerAsnGluAsnCys 371  
 Db 242 TCTTTAATGGCTACTGCTGTACAAATCCATGGAGCTGAGTGACATTTCCAATGAAAACCTGC 301  
 QY 372 ArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGlu 391  
 Db 302 CGAATAAACACATATGCTACTTTCAGAGCCACCATCACATTTCTAGAGGGGATCTCTGGAA 361  
 QY 392 ValAsnIleIleGlnValAlaAlaAspValProIleProThrLeuGlnProAspAsnSerLeu 411  
 Db 362 GTCAGCATCATGAGATAGACAGATGTCCCATGCCACACCCGAGCTGCCAACCTCCCTG 421  
 QY 412 MetAspPheIleValThrCysLysGlyValAlaThrProThrGluAlaCysThrIleIleSer 431  
 Db 422 ATGGACTTCACTGTGACCTGCAAAAGGGCCACCCCATGGAAGCCTGTACGATCATCTCC 481  
 QY 432 AspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeu 451  
 Db 482 GACCCACCTGCCAGATCGCCAGAACCGGTCTGCAAGCCTGTGGCTGTGGATGGCTG 541  
 QY 452 CysLeuLeuSerValArgAlaPheAsnGlySerGlyThrTyrCysValAsnPhaThr 471  
 Db 542 TGCTG-CTGTCTGTGAGAAGAGCCTTCAATGGGTCTGGCACCTACTGTGTGAATTTCACT 600  
 QY 472 LeuGlyAspAspAlaSer 477  
 Db 601 CTGGGAGATGATGCAAGC 618

RESULT 5

US-08-594-031-105/c  
 ; Sequence 105, Application US/08594031  
 ; Patent No. 5783182  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THOMPSON, Timothy C.  
 ; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
 ; NUMBER OF SEQUENCES: 175  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: BAKER & BOTTS, L.L.P.  
 ; STREET: 1299 Pennsylvania Avenue, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20004-2400  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq Version 1.5

;> CURRENT APPLICATION DATA:  
;> APPLICATION NUMBER: US/08/594,031  
;> FILING DATE: 30-JAN-1996  
;> CLASSIFICATION: 435  
;> PRIOR APPLICATION DATA:  
;> APPLICATION NUMBER: 60/006,838  
;> FILING DATE: 16-NOV-1995  
;> ATTORNEY/AGENT INFORMATION:  
;> NAME: Remenick, James  
;> REGISTRATION NUMBER: 36,902  
;> REFERENCE/DOCKET NUMBER: 0A146-0110  
;> TELECOMMUNICATION INFORMATION:  
;> TELEPHONE: 202-639-7700  
;> TELEFAX: 202-639-7890  
;> TELEX:  
;> INFORMATION FOR SEQ ID NO: 105:  
;> SEQUENCE CHARACTERISTICS:  
;> LENGTH: 524 base pairs  
;> TYPE: nucleic acid  
;> STRANDEDNESS: single  
;> TOPOLOGY: linear  
;> MOLECULE TYPE: cDNA  
;> HYPOTHETICAL: NO  
;> ANTI-SENSE: NO  
;> FRAGMENT TYPE:  
;> ORIGINAL SOURCE:  
;> US-08-594-031-105

Alignment Scores:  
Pred. No.: 4,22e-63 Length: 524  
Score: 729.00 Matches: 150  
Percent Similarity: 89.66% Conservative: 6  
Best Local Similarity: 86.21% Mismatches: 18  
Query Match: 23.82% Indels: 2  
DB: 1 Gaps: 0

US-09-943-075A-2 (1-572) x US-08-594-031-105 (1-524)

Qy 389 ILeuGluValAsnIleleGlnValAlaAspValProIleProThrLeuGlnProAsp 408  
Db 522 ATCTGGAAGTCAGCATCATGAGATAGCAGATGTCCTCCATGCCACACCGACCTGCC 463  
Qy 409 AsnSerLeuMetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCysThr 428  
Db 462 AACTCCTGATGACITTCATCTGACCTGCAAGGGGCCACCCCATGGAAGCCTGTACG 403  
Qy 429 IleIleSerAspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAlaVal 448  
Db 402 ATCATCTCCGACCCACCTGCGCAGATCGCCAGAACCGGGTCTGCAGCCCTGTGGCTGTG 343  
Qy 449 AspGluLeuCysLeuLeuSerValArgAlaPheAsnGlySerGlyThrTyrCysVal 468  
Db 342 GATGGCTGTGCTG -CTGTCCTGAGAGAGCCTTCATGGTGTCTGGCACCCTACTGTGTG 284  
Qy 469 AsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIlePro 488  
Db 283 AATTTCACTCTGGAGATGATCGACGATGCGCCCTCCACGACCCCTGATCTATATCCCT 224  
Qy 489 GlyLysAspLeuGlySerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCys 508  
Db 223 GGCAAGACCC -AGACTCCCTCTGAGAGCAGTGAATGGTGTCTCTATCCATCCATCGCTGC 165  
Qy 509 LeuAlaMetPheValThrMetValThrIleLeuLeuTyrLysLysHisLysThrTyrLys 528  
Db 164 CTGGCTGTGCTGTGTACCATGTGTACCATCTTGCTGTACAAAAACACAGGGCGTACAAG 105  
Qy 529 ProIleGlyAsnCysThrArgAsnValValLysGlyLysGlyLeuSerValPheLeuSer 548  
Db 104 CCAATAGGAACCTGCCCAAGAACACGCTCAAGGGCAGGGCCCTGAGTGTCTCTCTCACT 45  
Qy 549 HisAlaLysAlaProPheSerArgLysAspArgGluLysAsp 562  
Db 44 CACGCGAAAGCCCGCTTCTCCGAGGAGACACGAGAGGAT 3

## RESULT 6

US-08-594-031-115  
; Sequence 115, Application US/08594031  
; Patent No. 5783182  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/594,031  
; FILING DATE: 30-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: 0A146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700  
; TELEFAX: 202-639-7890  
; TELEX:

INFORMATION FOR SEQ ID NO: 115:

SEQUENCE CHARACTERISTICS:  
LENGTH: 473 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-594-031-115

Alignment Scores:  
Pred. No.: 5,47e-62 Length: 473  
Score: 717.00 Matches: 136  
Percent Similarity: 90.45% Conservative: 6  
Best Local Similarity: 86.62% Mismatches: 13  
Query Match: 23.42% Indels: 2  
DB: 1 Gaps: 1

US-09-943-075A-2 (1-572) x US-08-594-031-115 (1-473)

Qy 274 LeuAsnTyrSerAlaIleSerTyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheVal 293  
Db 2 CTCACGACCTCTGCCATTTCTTACAGTGAACCTTTGGGACACACACCTGTTTGTG 61  
Qy 294 SerAsnAsnHisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeu 313  
Db 62 TCCAACAAATCACACTTTGAATCACAATCTATGTGCTCAATGGAACCTTCAACCTTAACCTC 121  
Qy 314 ThrValGlnThrAlaValProGlyProCysProSerProThrProSerProSerProSer 331  
Db 122 ACCGTGCAAACTCAGTGGCCGGGCATGCCCTCCCTTCGCTTCGACTCCGCTCCA 181  
Qy 332 SerSerThrSerProSerProAlaSerProSerProSerProThrLeuSerThrProSerPro 351

```
Db 182 CCTTCAACTCGCCCTCACCTCCGCTCACCTCTGCCCCACATATATCAACACCTAGCCCC 241
Qy 352 SerLeuMetProThrGlyTyrLysSerMetGluLeuSerAspLeuSerHisGluAsnGly 371
Db 242 TCTTTAAATGCTTACCTGCTTCAAAATCCATGAGCTGAGTACATTTTCCAAATGAAACTGC 301
Qy 372 ArgIleAsnArgTyrGlyTyrPheArgAlaThrIlePheIleValAspGlyIleLeuGlu 391
Db 302 CGAATAAACAGATATGGCTACTCTCAGAGCCACCATCACATTTAGAGGGATCTCTGGA 361
Qy 392 ValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnProAspHisSerLeu 411
Db 362 GTACAGCATCATGATAGCATAGATGTCCCATGCCACACCGAGCTGCCAACTCCCTG 421
Qy 412 MetAspPheIleValThrCysLysGlyValAlaThrProThrGluAlaCysThr 428
Db 422 ATGGACTTCACTGTGACCTGCAAGAGGGCCACCCCATGGAAGCTGTAGC 472

RESULT 7
US-09-056-105-1
: Sequence 1, Application US/09056105
: Patent No. 6287369
: GENERAL INFORMATION:
: APPLICANT: KIPPS, THOMAS J.
: APPLICANT: WU, YUNCI
: TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
: FILE OF INVENTION: PROCESSING
: FILE REFERENCE: 233/221
: CURRENT APPLICATION NUMBER: US/09/056,105
: CURRENT FILING DATE: 1998-04-06
: EARLIER APPLICATION NUMBER: 60/043,467
: EARLIER FILING DATE: 1997-04-10
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2130
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-056-105-1

Alignment Scores:
Pred. No.: 4,55e-52 Length: 2130
Score: 627.50 Matches: 176
Percent Similarity: 41.85% Conservative: 96
Best Local Similarity: 27.08% Mismatches: 205
Query Match: 20.50% Indels: 173
DB: 4 Gaps: 19

US-09-943-075A-2 (1-572) x US-09-056-105-1 (1-2130)
Qy 7 ValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGlnAlaAlaLys-----Arg 24
Db 46 CTTCTTCATTGGCTGTGATAGTGTCTGCTGCTGGGGGTACAAAGTACCACGA 105
Qy 25 PheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGluAsnGlnLeu 44
Db 106 AACAGGAGTGGCTGGT-----GTCTCAAGGCAACTC 138
Qy 45 ArgGlyTyrSerSerAspGluAsnGluTrpAspGluLeuTyrProValTrpArg 64
Db 139 AGA-----ACCAAGCCTGGAACAGGAGCGCTGTATCCAGAGTGGACA--- 180
Qy 65 GlyGluGlyArgTyrLysAspSerTrpGluGlyValGlnAlaAlaLeuThrSer 84
Db 181 ---GAAGCCACAGACTGACTGCTGGAGAGGTGGTCAAGTGTCCCTCAAGGTCAAGTAAT 237
Qy 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeuValPheProArg 104
Db 238 GATGGCCCTACACTGATGGTGAATATGCTCTCTATTGCTTGAACCTTCCCTGGA 297
Qy 105 CysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsn---CysArgSerAsp 123
Db 298 AGCCAAAAGGTATTGCCAGATGGCAGGTATTCTGGGTCAACAATACCATCATCATGGG 357
```

```
Qy 124 LeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAspAspGluAsp 143
Db 358 AGCAGGTGTGGGAGGACAGCCAGTGTATCCCCAGGAAACT-----GACGATGCC--- 408
Qy 144 TrpGluAspAsnThrSerGlnGlnHisLeuArgPheProAspGlyLysPtoPhePro 163
Db 409 -----TGCACTCTCCCTGATGGTGCACCTTGCCTCCCA 438
Qy 164 ArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeuGlyGlnTyr 183
Db 439 TCTGGCTCTTGGTCTCAGAAGAGAGCTTTGTTATGTCTGGAAGACTGGGCCAATAC 498
Qy 184 PheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsnLeuThrVal 203
Db 499 TGGCAAGTTCTAGGGGGCCAGTGTCTGGGTGACATTTGGACAGCAGCGGCAATGCTG 558
Qy 204 GlyProGlnValMetGluValIleValPheArgArgHisGly---ArgAlaIleTyrIlePro 222
Db 559 GGCACACACCATGGAAGTGACTGTCTACCATGCGCGGGGATCCCGAGCTATGTGCCT 618
Qy 223 IleSerLysValIlyAspValTyrValIleThrAspGlnIlePheIlePheValThrMet 242
Db 619 CTTGCTCATTTCCAGCTCAGCCTTACCATTTACTGACAGGTGCTCTTCTCCGTGAGGCTG 678
Qy 243 TyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhe 262
Db 679 TCCAGTTTGGGGCCCTTGGATGGAGGAAACAAGCATTCTCTGAGAAATCAGCCTCTGACC 738
Qy 263 PheAspValIleIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLys 282
Db 739 TTTGGCCTCCAGCTCCATGACCCCACTGCTATCTGGCTGAAGCTGACCTCTCTCTACACC 798
Qy 283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302
Db 799 TGGGACTTTGGAGACAGTAGTGGAAACCTGTCTCTCGGCACACTTGTGGTCACTCATACT 858
Qy 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro----- 320
Db 859 TACCTGGAGCCTGGCCCACTGCTCCCAAGTGTGCTGCTGAGGCTGCCATTCTCTCTCACC 918
Qy 321 -----GlyProCysPro----- 324
Db 919 TCTGTGGCTCTCTCCCACTGCTCCAGGCACACAGATGGGCACAGGCCAAGTGGAGAGGCC 978
Qy 324 ----- 324
Db 979 CCTAACACACAGCTGGCCCAAGTGCCTACTACAGAAGTTGTGGTACTACACCTGGTCTAG 1038
Qy 325 SerProThrProSerProSerSerSerThr----- 334
Db 1039 GCCTCAACTGCGAGAGCCCTCTGGAAACCACTCTGTGCAGGTGGCCAACTGAGTCAATA 1098
Qy 335 SerProSerProAlaSerSerProSerProThrLeuSerThrProSerProSerLeuMet 354
Db 1099 AGCACTGCACCTGTGTCAGATGCCAACTCAGAGAGCAGCAGGTATGACACCTGAGAAGGTG 1158
Qy 355 Pro-----ThrGly 357
Db 1159 CCATTTTCAGAGGTGATGGGTACCACACTGGCAGAGATGTCAACTCCAGAGGCTACAGGT 1218
Qy 358 TyrLysSerMetGluLeuSerAspIle----- 366
Db 1219 ATGACACCTTGCAGAGGTATCAATTGTGTGCTTTCTGSAACCAACAGCTGCACAGGTAACA 1278
Qy 366 ----- 366
Db 1279 ACTACAGAGTGGGTGGAGACACACAGCTAGAGAGTACCTATCCCTGAGAGCTGAAGTCCA 1338
Qy 367 -----SerAsnGlu----- 369
Db 1339 GATGCCAGCTCAATCATGTCTACGGAAGATATTACAGGTTCCCTGGGCCCTCTGCTGGAT 1398
```

```
Qy 370 -----AsnCysArgIleAsnArg 375
Db 1399 GGTACAGCCACCTTAAGGCTGTGAAGACAGACCAAGTCCCGCTGGATTGTCTGTATCGA 1458
Qy 376 TyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle 395
Db 1459 TATGGTTCCTTTCCTCACCCTGCACATTGTCCAGGATTAAGAGTCCGAGATCCCTG 1518
Qy 396 GluValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
Db 1519 CAGGCT-----GTGCCGTGCGGTGAGGGGAT-----GCATTGAGCTGACT 1560
Qy 416 ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435
Db 1561 GTGTCTCCAGAGCGGTGCTCCCAAGGAAGCCTGCATGAGATCTCATCGCCAGGGTGC 1620
Qy 436 GluIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSer 455
Db 1621 CAGCCCTCCCTGCCAGCGGTGTGCCAGCCTGTCTACCCAGCCAGCCTGCCAGCTGGTT 1680
Qy 456 ValArgArgAlaPhe---AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAsp 474
Db 1681 CTGCACAGACTACTGAGGCTGCTCGGGGACACTACTGCCATATGTCTCTGGCTGAT 1740
Qy 475 AspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGly--- 493
Db 1741 ACCAACAGCCTGGCAGTGTGCACAGCCAGCCTTATCATGCTGTCAAGAAGCAGGCTT 1800
Qy 494 SerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheVal 513
Db 1801 GGCAGGTTCCTGATCTGTGGGCACTGTG-----CTGGTGTGATGGCT 1845
RESULT 8
US-08-417-174-26
; Sequence 26, Application US/0841174
; Patent No. 584075
; GENERAL INFORMATION:
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
; APPLICANT: STEVEN A.
; TITLE OF INVENTION: MELANOMA ANTIGENS AND
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
; NUMBER OF SEQUENCES: 126
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,174
; FILING DATE: 05-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,565
; FILING DATE: 22-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4124US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
```

```
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172
; TYPE: nucleotide
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: cDNA
; US-08-417-174-26
Alignment Scores:
Pred. No.: 5,27e-52 Length: 2172
Score: 627.00 Matches: 174
Percent Similarity: 41.03% Conservative: 96
Best Local Similarity: 26.44% Mismatches: 206
Query Match: 20.48% Indels: 182
DB: Gaps: 19
US-09-943-075A-2 (1-572) x US-08-417-174-26 (1-2172)
Qy 17 LeuProLeuGlnAlaAlaLysArgPheArgAspVal----- 28
Db 12 TTACCAATCCCGACCGGAGAACACATGATCTGGTCTAAAAAGATGCCCTTCTTCAT 71
Qy 29 -----LeuGlyHisGluGlnTyrPro-----Asp 36
Db 72 TTGGCTGTGATAGTCTTTTGGCTGTGGGGCTACAAAAGTACCCAGAAACAGGAC 131
Qy 37 HisMetArgGluAsnAsnGlnLeuArgGlyTyrSerSerAspGluAsnGluTyrAspGlu 56
Db 132 TGGCTGTGGTCTCAAGCAACTCAGA-----ACCAAGCCTGGAAACAGG 176
Qy 57 GlnLeuTyrProValTrrpArgArgGlyGluGlyArgTrrpLysAspSerTrpGluGlyGly 76
Db 177 CAGCTGTATCCAGAGTGGACA-----GAAGCCAGAGACTTGACTGCTGGAGAGTGGT 230
Qy 77 ArgValGlnAlaAlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPhe 96
Db 231 CAAGTGTCCCTCAAGGTCAAGTAATGATGGCTACACTGATTTGGTGCNAATGCCCTCCTC 290
Qy 97 ValValAsnLeuValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyr 116
Db 291 TCTATTGCCCTTGAACCTCCCTCGAAGCCAAAGGTATTGCCAGATGGCAGGTATCTGG 350
Qy 117 GluArgAsn---CysArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrp 135
Db 351 GTCACCAATACCATCATCATGGAGCCAGGTGTGGGGAGGACAGCCAGTGTATCCCCAG 410
Qy 136 ThrThrGlyAlaAspAspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArg 155
Db 411 GAAACT-----GACGATGCC-----TGCATC 431
Qy 156 PheProAspGlyLysProPheProArgProHisGlyArgLysLysTrpAsnPheValTyr 175
Db 432 TTCCTGTAGTGTGACCTTCCCATCTGCTGCTCTCAGAGAGAGAGCTTTGTTAT 491
Qy 176 ValPheHisThrLeuGlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSer 195
Db 492 GTCTGGAAGACCTGGGGCCAAATACTGGCAATTTCTAGGGGGCCCGAGTGTCTGGGCTGAGC 551
Qy 196 IleAsnThrValAsnLeuThrValGlyProGlnValMetGluValIleValPheArgArg 215
Db 552 ATTGGGACAGCGCAGGCAATGCTGGGCACACACACCATGGAGTGCATGCTACCATCGC 611
Qy 216 HisGly---ArgAlaTyrIleProIleSerLysValLysAspValTyrValIleThrAsp 234
Db 612 CGGGGATCCCGGAGCTATGTGCTCTTCCAGCTCAGCTTCCAGCTTCCACATTTACTGAC 671
Qy 235 GlnIleProIlePheValThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThr 254
Db 672 CAGGTGCTTTCTCCGTGAGCGTGTCCCAAGTTGCCGGGCTTGGATGGAGGGAACAAGCAC 731
Qy 255 PheLeuArgAspLeuProIlePheAspValLeuIleHisAspProSerHisPheLeu 274
```



```

177 CAGCTGTATCCAGAGTGGACA-----GAAGCCACAGACTTGACTGCTGGAGAGTGGT 230
Qy 77 ArgValGlnAlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPhe 96
Db 231 CAAAGTGCCTCAAGGTCAATGATGAGGCTACACTGATGGTGGCAATGCTCTTC 290
Qy 97 ValValAsnLeuValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyr 116
Db 291 TCTATTGCCCTTGAACCTCCCTGGAAGCCAAAGGTAFTTGCAGATGGCAGGTTATCTCG 350
Qy 117 GluArgAsn---CysArgSerAspLeuLeuAlaSerAspProTyrValTyrAsnTrp 135
Db 351 GTCACAAATACCATCATCAATGGAGCCAGGTTGGGAGGACAGCCAGGTATPCCCCAG 410
Qy 136 ThrThreGlyAlaAspAspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArg 155
Db 411 GAAACT-----CACGATGCC-----TGCAATC 431
Qy 156 PheProAspGlyLysProPheProArgProHisGlyArgLysLysTyrAsnPheValTyr 175
Db 432 TTCCTGATGGTGGACCTTGCCTCTGCTCTGCTCTCAGAGAGAACTTGTGTTAT 491
Qy 176 ValPheHisThrLeuGlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSer 195
Db 492 GTCTGGAAGACCTGGGGCCCACTGGCAATTTCTAGGGGGCCAGCTGTCTGGGCTGAGC 551
Qy 196 IleAsnThrValIleAsnLeuThrValGlyProGlnValMetGluValIleValPheArgArg 215
Db 552 ATGGGACAGCGGAGGCAATGTGGGCACACACACACCAATGGAGTACTGTCTACCATCGC 611
Qy 216 HisGly---ArgAlaTyrIleProIleSerLysValLysAspValTyrValIleThrAsp 234
Db 612 CGGGGATCCCGAGCTATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 671
Qy 235 GlnIleProIlePheValThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThr 254
Db 672 CAGGTGCTTTCCTGAGCGGTGCCAGTTCGGGGCTTGGATGGAGGGAACACAGCAC 731
Qy 255 PheLeuArgAspLeuProIlePhePheAspValLeuIleHisAspProSerHisPheLeu 274
Db 732 TTCCTGAGAAATCAGCTCTGACCTTGCCTTCCAGCTCCATGACCCAGCTGCTATCTG 791
Qy 275 AsnTyrSerAlaIleSerTyrIleTyrPheGlyAspAsnThrGlyLeuPheValSer 294
Db 792 GCTGAAGCTGACCTCTCTACACTGGGACTTTGGAGACAGTACTGGAAACCTGATCTCT 851
Qy 295 AsnAsnHisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThr 314
Db 852 CGGGCAGCTGTGTCTACTATCTTACCTGGAGCGCTGGCCAGCTCACTGCCAGGTGGTC 911
Qy 315 ValGlnThrAlaValPro-----GlyProCysPro----- 324
Db 912 CTGCAGGCTGCCATTCCTCTCACTCCTCTGCTCTCTCCAGTTCACAGGCACACAGAT 971
Qy 324 ----- 324
Db 972 GGGCAGCGCCAACTGCAGAGGCGCCCTAACACACAGCTGGCCAAAGTGCCTTACTACAGAA 1031
Qy 325 -----SerProThrProSerProSerSerSerSerSerSerSerSerSerSerSer 334
Db 1032 GTTGTGGGTACTACACCTGGTCAGCGCCCACTGCAGAGCCCTCTGGAAACACATCTGTG 1091
Qy 335 -----SerProSerProAlaSerSerProSerProSerProSerProSerProSer 346
Db 1092 CAGGTGCCAACCACTGAAGTCAATAGACACTGCACCTGTGCAGATGCCCACTGCAGAGAGC 1151
Qy 347 SerThrProSerProSerLeuMetPro----- 355
Db 1152 ACAGGTATGACACCTCAGAGAGGTGCCAGTTTCAGAGGTCAATGGGTACACACATGGCAGAG 1211
Qy 356 -----ThrGlyTyrLysSerMetGluLeuSerAspIle----- 366

```

```

Db 1212 ATGTCAACTCCAGAGCTACAGGTATGACACCTGCAGAGGTATCAATTTGTGGTCTTCT 1271
Qy 366 ----- 366
Db 1272 GGAACACACACTGCACAGGTACAACTACAGAGTGGTGGAGACCACAGCTAGAGAGCTA 1331
Qy 367 -----SerAsnGlu----- 369
Db 1332 CCTATCCTCAGCCTGAAGGTCCAGATGCCAGTCAATCATCTCTACGGAAGATTATACA 1391
Qy 369 ----- 369
Db 1392 GGTTCCTCGGGCCCCCTGCTGGATGGTACAGCACCTTTAAGCTGGTGAACAGACAGTC 1451
Qy 370 -----AsnCysArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAsp 387
Db 1452 CCCCTGGATGTGTCTGTATGATGATGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1511
Qy 388 GlyIleLeuGluValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnPro 407
Db 1512 GGTATCGAAAGTCCGAGATCTCTGAGGCT-----GTCCCTCGCTGAGGGG 1559
Qy 408 AspAsnSerLeuMetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCys 427
Db 1560 GAT-----GCATTTGAGCTGACTGTCTCTGCCAAGGGGGCTGCCCAAGAAAGCCTGC 1613
Qy 428 ThrIleIleSerAspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAla 447
Db 1614 ATGGAGATCTCATCGCCAGGGTGCACAGCCCCCTGCCAGGGCTGTGCCAGCTGTGCTA 1673
Qy 448 ValAspGluLeuCysLeuLeuSerValArgArgAlaPhe---AsnGlySerGlyThrTyr 466
Db 1674 CCCAGCCACCTCCAGCTGCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1733
Qy 467 CysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuIleSer 486
Db 1734 TGCTCTCAATGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793
Qy 487 IleProGlyLysAspLeuGly---SerProLeuArgThrValAsnGlyValLeuIleSer 505
Db 1794 ATGGCTGGTCAAGAACAGGCGCTTGGCAGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 1847
Qy 506 IleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeuTyrLysLys 523
Db 1848 -----CTGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1892

```

## RESULT 10

```

US-09-007-961-26
; Sequence 26, Application US/09007961
; Patent No. 5994523
; GENERAL INFORMATION:
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
; APPLICANT: STEVEN A.
; TITLE OF INVENTION: MELANOMA ANTIGENS AND
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007,961
; FILING DATE:
; CLASSIFICATION:

```



: PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: 08/231,565  
 : FILING DATE: 22-APR-1994  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: CAROL M. GRUPPI  
 : REGISTRATION NUMBER: 37,341  
 : REFERENCE/DOCKET NUMBER: 2026-4124  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (212) 758-4800  
 : TELEFAX: (212) 751-6849  
 : TELEX: 421792  
 : INFORMATION FOR SEQ ID NO: 26:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2172  
 : TYPE: nucleotide  
 : STRANDEDNESS: Double  
 : TOPOLOGY: Unknown  
 : MOLECULE TYPE: cDNA  
 : US-09-007-961-26

Alignment Scores:  
 Pred. No.: 5,27e-52 Length: 2172  
 Score: 627.00 Matches: 174  
 Percent Similarity: 41.03% Conservative: 96  
 Best Local Similarity: 26.44% Mismatches: 206  
 Query Match: 20.48% Indels: 182  
 DB: Gaps: 19

US-09-943-075a-2 (1-572) x US-09-007-961-26 (1-2172)

Qy	216	HisGly---ArgAlaTyrIleProIleSerLysValLysAspValTyrValIleThrAsp	234
Db	612	CGGGATCCCGGAGCTATGCTCTCTGCTCATTTCCAGCTCAGCTTCCACATTTACTGAC	671
Qy	235	GlnIleProIlePheValThrMetTyrGlnLysAsnAspArgAsnSerSerGluThr	254
Db	672	CAGTGCCTTTCTCCGTCAGCGTCCCGATGTCGGGGCTTGGATGCGAGGAAAGACAC	731
Qy	255	PheLeuArgAspLeuProIlePhePheAspValLeuIleHisAspProSerHisPheLeu	274
Db	732	TTCTGAGAAATCAGCTCTGACCTTGCCTCCAGCTCCATGACCCCGGCTGCTCTG	791
Qy	275	AsnTyrSerAlaIleSerTyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSer	294
Db	792	GCTAAAGCTGACCTCTCTACACCTGGGAGCTTTGGAGACAGTACTGGAACCTGATCTCT	851
Qy	295	AsnAsnHisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThr	314
Db	852	CGGTCACCTTGTGTGCTACTCATCTTACCTGGAGCCTGGCCCGCTACTGCCAGGTGTC	911
Qy	315	ValIleThrAlaValPro-----GlyProCysPro-----	324
Db	912	CTGCGAGGCTGCCATTCCTCTCCTCCTGCTGCTCCTCCAGTTCCAGGCACACAGAT	971
Qy	324	-----	324
Db	972	GGGACAGGCCAACTGCAGAGGCCCTAACCCACAGCTGGCCAGTGCCTACTACAGAA	1031
Qy	325	-----SerProThrProSerProSerSerThr-----	334
Db	1032	GTTCGTGGTACTACACCTTGTGTCAGCGCCCACTGCAGAGCCTCTGGAACCATCTGTG	1091
Qy	335	-----SerProSerProAlaSerSerProSerProThrLeu	346
Db	1092	CAGTGGCAACCACTGAAGTCAATAGCACTGCNCCTGTGCAGATGCCAATGTCAGAGGC	1151
Qy	347	SerThrProSerProSerLeuMetPro-----	355
Db	1152	ACAGGTATGACACCTCAGAAAGTGCAGGTTTCAGAGGTCATGGGTACCACACTGGCAGAG	1211
Qy	356	-----ThrGlyTyrLysSerMetGluLeuSerAspIle-----	366
Db	1212	ATGTCAACTCCAGAGCTACAGGTATGACACCTGCAGAGGTATCAATGTGGTCTTCT	1271
Qy	366	-----	366
Db	1272	GGAAACACAGCTGCACAGGTAACAACTACAGAGTGGGTGGAGACCACAGCTAGAGACTA	1331
Qy	367	-----SerAsnGlu-----	369
Db	1332	CCTATCCCTGAGCCTGAAGTCCAGTGCACCTCAATCATGTCTACGGAAAGTATTACA	1391
Qy	369	-----	369
Db	1392	GGTTTCCTGGGCCCTGCTGGATGGTACAGCCACCTTAAGGCTGGTGAAGAGACAAGTC	1451
Qy	370	-----AsnCysArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAsp	387
Db	1452	CCCTCGATTGTCTGTGTATGATATGTTCTTTCCTCCGTCACCCCTGGACATTTGTCAG	1511
Qy	388	GlyIleLeuGluValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnPro	407
Db	1512	GGTATTTGAAAGTGCAGAGATCCTGAGGCT-----GTGGCGTCCGGTGGAGGG	1559
Qy	408	AspAsnSerLeuMetAspPheIleValThrCysLysGlyValAlaThrProThrGluAlaCys	427
Db	1560	GAT-----GCATTTGAGCTGACTGTGCTCTCCAAAGGGGGCTGCCCAAGGAAGCTGC	1613
Qy	428	ThrIleIleSerAspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAla	447
Db	1614	ATGGAGATCTCATCGCAGGGTGCAGGCCCTGTCGCCAGCGCTGTGCCAGCTTGCTA	1673

Qy	17	LeuProLeuGlnAlaAlaLysArgPheArgAspVal-----	28
Db	12	TTACCAATCCGCGGGAAGACAACTGGATGCTGGTGTAAAAGATGCCCTTCAT	71
Qy	29	-----LeuGlyHisGluGlnTyrPro-----	36
Db	72	TTGGCTGTGATAGTGTCTTGTGCTGTGGGGCTACAAAAGTACCCAGAACCCAGGAC	131
Qy	37	HisMetArgGluAsnAsnGlnLeuArgGlyTrpSerSerGluAsnGluTrpAspGlu	56
Db	132	TGGCTGTGCTCTCAAGGCAACTCAGA-----ACCAAGCTCGAAGCAGG	176
Qy	57	GlnTyrProValTrpArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGly	76
Db	177	CAGCTGATCCAGTGGACA-----GAAGCCAGAGACTTGACTGCTGGAGAGTGGT	230
Qy	77	ArgValGlnAlaAlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPhe	96
Db	231	CAAGTGTCCCTCAAGCTCAGTAAATGATGGCCCTACACTGATTGTGCAATGCTCCTTC	290
Qy	97	ValValAsnLeuValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyr	116
Db	291	TCTATTCCTGTGAACTTCCTCGAGCCAAAAGTATTGCCAGATGGCGAGTTATCTGG	350
Qy	117	GluArgAsn---CysArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrp	135
Db	351	GTCAACAATACCATCATCATGAGCCAGGAGTGTGGGAGGACAGCCAGTGTATCCOCAG	410
Qy	136	ThrThrGlyAlaAspAspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArg	155
Db	411	GAAACT-----GACGATGCC-----TGCATC	431
Qy	156	PheProAspGlyLysProPheProArgProHisGlyArgLysLysTrpAsnPheValTyr	175
Db	432	TTCCCTGATGCTGGACCTTGGCCATCTGGCTCTTGTGCTCAGAGAGAAGCTTTGTTAT	491
Qy	176	ValPheHisThrLeuGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSer	195
Db	492	GTCTGGAGACCTCGGGCCAACTAGGCAATTTCTTAGGGGGCCCGAGTGTCTGGCTGAGG	551
Qy	196	IleAsnThrValAsnLeuThrValGlyProGlnValMetGluValIleValPheArgArg	215
Db	552	ATTGGGACAGGCAAGCAATGCTGGGCAACACACCATGGAAGTGACTGTCTACCATCGC	611

QY 448 ValAspGluLeuCysLeuLeuSerValArgAlaPhe---AsnGlySerGlyThrTyr 466  
 Db 1674 CCCAGCCAGCGTCCGACCTGGTCTCCACAGATACACTGAAGGTTGGCTCGGGGACATAC 1733  
 QY 467 CysValAsnPheThrLeuGlyAspAlaSerLeuAlaLeuThrSerAlaLeuLeuSer 486  
 Db 1734 TGCCTCAATGTCTCTGGCTGATACCAAGAGCTGGCAGTGGTCCAGCACCAGCTTATC 1793  
 QY 487 IleProGlyLysAspLeuGly---SerProLeuArgThrValAsnGlyValLeuLeuSer 505  
 Db 1794 ATGCCTGGTCAAGAGCAGCGCTTGGCAGGTTCCGCTGATCTGGGCACTTG----- 1847  
 QY 506 IleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeuTyrLys 523  
 Db 1848 -----CTGGTGTGATGGTGTGGTCTTGCATCTCTGATATATAGGCGC 1892

RESULT 11  
 ; US-09-267-439-26  
 ; Sequence 26, Application US/09267439  
 ; Patent No. 6270778  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
 ; APPLICANT: STEVEN A.  
 ; TITLE OF INVENTION: MELANOMA ANTIGENS AND  
 ; THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
 ; METHODS  
 ; NUMBER OF INVENTION: 126  
 ; NUMBER OF SEQUENCES: 126  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
 ; STREET: 345 PARK AVENUE  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: USA  
 ; ZIP: 10154

COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY DISK  
 COMPUTER: IBM PC COMPATIBLE  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/267,439  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/417,174  
 FILING DATE: 05-APR-1995  
 APPLICATION NUMBER: US/08/231,565  
 FILING DATE: 22-APR-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CAROL M. GRUPPI  
 REGISTRATION NUMBER: 37,341  
 REFERENCE/DOCKET NUMBER: 2026-4124US1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 758-4800  
 TELEFAX: (212) 751-6849  
 TELEX: 421792  
 INFORMATION FOR SEQ ID NO: 26:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2172  
 TYPE: nucleotide  
 STRANDEDNESS: Double  
 TOPOLOGY: unknown  
 MOLECULE TYPE: cDNA  
 US-09-267-439-26

Alignment Scores:  
 Pred. No.: 5,27e-52 Length: 2172  
 Score: 627.00 Matches: 174  
 Percent Similarity: 41.03% Conservative: 96  
 Best Local Similarity: 26.44% Mismatches: 206  
 Query Match: 20.48% Indels: 182  
 DB: 4 Gaps: 19

US-09-943-075A-2 (1-572) x US-09-267-439-26 (1-2172)

QY 17 LeuProLeuGlnAlaAlaLysArgPheArgAspVal----- 28  
 Db 12 TTACCAATCGCGACCGGGAAGAACACACATGATCTGGTGTCTAAAGATGCCTTCTTCAT 71  
 QY 29 -----LeuGlyHisGluGlnTyrPro-----Asp 36  
 Db 72 TTGGCTGTGATAGGTGCTTTTGGCTGTGGGGCTCAAAAGATTACCCAGAAACACAGGAC 131  
 QY 37 HisMetArgGluAsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGlu 56  
 Db 132 TGGCTGTGGTGTCTCAAGCAACTCAGA-----ACCAAGGCTCGAACAGG 176  
 QY 57 GlnLeuTyrProValTrpArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGly 76  
 Db 177 CAGCTGTATCCAGAGTGGACA-----GAAGCCAGAGACTTACTGCTGGAGAGTGT 230  
 QY 77 ArgValGlnAlaAlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPhe 96  
 Db 231 CAAGTGTCCCTCAAGTCAAGTCAATGATGGGCTACACTGATGGTCAATGCTCCTTC 290  
 QY 97 ValValAsnLeuValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyr 116  
 Db 291 TCTATTGCTTGAACCTTCCCTGGAAAGCCAAAGGTATTGCCAGATGGCAGGTTATCTGG 350  
 QY 117 GluArgAsn---CysArgSerLeuLeuAlaSerAspProTyrValTyrAsnTrp 135  
 Db 351 GTCACATATACCATCATCAATGGGAGCCAGGTGTGGGAGACAGCCAGTATCCCCAG 410  
 QY 136 ThrThrGlyAlaAspAspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArg 155  
 Db 411 GAAACT-----GACGATGCC-----TGCATC 431  
 QY 156 PheProAspGlyLysProPheProArgProHisGlyArgLysLysTrpAsnPheValTyr 175  
 Db 432 TTCCTGTATGGTGGACCTGCCATCTGGCTTGGTCTCAGAAGAGAACTTTGTTAT 491  
 QY 176 ValPheHisThrLeuGlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSer 195  
 Db 492 GTCTGGAAGACCTGGGGCAATCTGGCAATTTCTAGGGGGCCCGTGTGGGCTGAGC 551  
 QY 196 IleAsnThrValAsnLeuThrValGlyProGlnValMetGluValIleValPheArgArg 215  
 Db 552 ATGGGACAGCAGCGGCAATGCTGGGCACACACACCATGGAAGTGAAGTGTCTACCATCGC 611  
 QY 216 HisGly---ArgAlaTyrIleProIleSerLysValLysAspValTyrValIleThrAsp 234  
 Db 612 CGGGATCCCGAGCTATGTGCTCTTCTCATTTCCAGCTCAGCTTCCACCATTTACTGAC 671  
 QY 235 GlnIleProIlePheValThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThr 254  
 Db 672 CAGGTGCCCTTCTCCGTGAGCGTGTCCAGTGTGGGGCTTGGAGGAGAAACAAGCAC 731  
 QY 255 PheLeuArgAspLeuProIlePhePheAspValLeuIleHisAspProSerHisPheLeu 274  
 Db 732 TTCCTGAGAAATCAGCCTCTGACCTTTCCTCCAGTCCATGACCCAGCTGCTATCTG 791  
 QY 275 AsnTyrSerAlaIleSerTyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSer 294  
 Db 792 GCTGAAGCTGACCTCTCTACACCTGGGACTTTGGAGACAGTAGTGGAAACCTGATCTCT 851  
 QY 295 AsnAsnHisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThr 314  
 Db 852 CGGGCACTTGTGGTCACTACATCTTACCTGGAGCCTGGCCAGCTGAGTCCAGGTTGTC 911  
 QY 315 ValGlnThrAlaValPro-----GlyProCysPro----- 324  
 Db 912 CTGCAGGCTGCCATTCCTCTCACCTCTGCTGCTCTCCCTCCAGGTTCCAGGACACAGAT 971  
 QY 324 ----- 324  
 Db 972 GGGCACAGGCCAACTGCGAGAGGCCCTTAACACACAGCTGGCCAAAGTGCCTACTACAGAA 1031



GENERAL INFORMATION:  
APPLICANT: THOMPSON, Timothy C.  
TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
NUMBER OF SEQUENCES: 175  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BAKER & BOTTS, L.L.P.  
STREET: 1299 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004-2400  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/594,031  
FILING DATE: 30-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/006,838  
FILING DATE: 16-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Remenick, James  
REGISTRATION NUMBER: 36,902  
REFERENCE/DOCKET NUMBER: 0A146-0110  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-639-7700  
TELEFAX: 202-639-7890  
TELEX:

## INFORMATION FOR SEQ ID NO: 94:

## SEQUENCE CHARACTERISTICS:

LENGTH: 335 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

US-08-594-031-94

## Alignment Scores:

Pred. No.:	3,64e-43	Length:	335
Score:	524.00	Matches:	99
Percent Similarity:	88.70%	Conservative:	3
Best Local Similarity:	86.09%	Mismatches:	9
Query Match:	17.12%	Indels:	4
DB:	1	Gaps:	1

US-09-943-075a-2 (1-572) x US-08-594-031-94 (1-335)

Qy	225	LysValLysAspValTyrValIleThrAspGlnIleProIlePheValThrMetTyrGln	244
Db	1	AAGTCAAGAGATGTGTATGTGATAACAGATCAGATCCCTGTTATTCGTGACCATGTCCAG	60
Qy	245	LysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhePheAsp	264
Db	61	AAGAATGACAGGAAGTGTCTGATGAGATCTCTCAGAGACCTCCCATTCGTTTCGAT	120
Qy	265	ValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLysTrpAsn	284
Db	121	GTCTCATTCATGATCCAGCCACTTCCTCAACGACTCTGCCATTCTCAAGTGAAC	180
Qy	285	PheGlyAspAsnThrGlyLeuPheValIleSerAsnAsnHisThrLeuAsnHisThrTrpVal	304
Db	181	TTTGGGGACACACATGGCTGTGTCTCCACCAATCACACTTTGAATCACACTTAATG	240
Qy	305	LeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValProGlyProCysPro	324
Db	241	CTAATGGAACCTTCAACCTTAACCTCACCCTGCAAACTGCAGTGCCTGGCGGCATGCCCT	300

Qy 325 SerProThrProSerProSerSerThrSerProSerProAla 339  
Db 301 CCCCTTCGCT-----TCGACTCGGCTCCACCTCG 333  
RESULT 14  
US-08-594-031-96  
; Sequence 96, Application US/08594031  
; Patent No. 5783182  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/594,031  
; FILING DATE: 30-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: 0A146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700  
; TELEFAX: 202-639-7890  
; TELEX:

## INFORMATION FOR SEQ ID NO: 96:

## SEQUENCE CHARACTERISTICS:

LENGTH: 335 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

US-08-594-031-96

## Alignment Scores:

Pred. No.:	3,64e-43	Length:	335
Score:	524.00	Matches:	99
Percent Similarity:	88.70%	Conservative:	3
Best Local Similarity:	86.09%	Mismatches:	9
Query Match:	17.12%	Indels:	4
DB:	1	Gaps:	1

US-09-943-075a-2 (1-572) x US-08-594-031-96 (1-335)

Qy	225	LysValLysAspValTyrValIleThrAspGlnIleProIlePheValThrMetTyrGln	244
Db	1	AAGTCAAGAGATGTGTATGTGATAACAGATCAGATCCCTGTTATTCGTGACCATGTCCAG	60
Qy	245	LysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhePheAsp	264
Db	61	AAGAATGACAGGAAGTGTCTGATGAGATCTCTCAGAGACCTCCCATTCGTTTCGAT	120
Qy	265	ValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLysTrpAsn	284

Db 121 GTCCTCATTCATGCCACCCACTTCTCTCAAGCACTCGCCATTCTCTCAAGTGGAC 180  
QY 285 PheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThrTyVal 304  
Db 181 TTTGGGGACACACTGGCCCTGTTGTCTCCACAATCACACTTTGAATCACACTTATGTG 240  
QY 305 LeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValProGlyProCysPro 324  
Db 241 CTCATGGAACCTTCAACCTTACCTCACCGTGCAACCTGACAGTCCCGGGCCATGCCCT 300  
QY 325 SerProThrProSerProSerSerThrSerThrSerProSerProAla 339  
Db 301 CCCCTTCGGCT-----TCGACTCCGCCTCCACCTCG 333

## RESULT 15

US-08-594-031-123  
; Sequence 123, Application US/08594031  
; Patent No. 5783182  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/594,031  
; FILING DATE: 30-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: 0A146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700  
; TELEFAX: 202-639-7890  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 123:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 335 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
US-08-594-031-123

Alignment Scores:  
Pred. NO.: 3.64e-43 Length: 335  
Score: 524.00 Matches: 99  
Percent Similarity: 88.70% Conservative: 3  
Best Local Similarity: 86.09% Mismatches: 9  
Query Match: 17.12% Indels: 4  
DB: 1 Gaps: 1

US-09-943-075A-2 (1-572) x US-08-594-031-123 (1-335)

QY 225 LysValLysAspValThrValIleThrAspGlnIleProIlePheValThrMetTyGln 244

Db 1 AAAGTGAAAGATGTGTATGTGATAACAGATCAGATCCCTGATTCGTGACCATGCCCCAG 60  
QY 245 LysAsnAspA-gAsnSerSerAspGlyThrPheLeuArgAspLeuProIlePhePheAsp 264  
Db 61 AAGAATGACAGAACTTGTCTGTGATGAGATCTTCTCAGAGACCTCCCATCGTCTTCGAT 120  
QY 265 ValLeuIleHisAspProSerHisPheLeuAsnTySerAlaIleSerTyIleTrpAsn 284  
Db 121 GTCCCTCATTCATGATGCCAGCCACTTCTCAACGACTCTGCCATTCTTACAAGTGAAC 180  
QY 285 PheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThrTyVal 304  
Db 181 TTTGGGGACAAACACTGGCCTGTTGTCTCCACAACATCACACTTTGAATCACACTTATGTG 240  
QY 305 LeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValProGlyProCysPro 324  
Db 241 CTCATGGAACCTTCAACCTTAACTTACCTCACCGTGCAACCTGACAGTCCCGGGCCATGCCCT 300  
QY 325 SerProThrProSerProSerSerThrSerThrSerProSerProAla 339  
Db 301 CCCCTTCGGCT-----TCGACTCCGCCTCCACCTCG 333

Search completed: April 9, 2003, 13:08:30  
Job time : 84 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run On: April 9, 2003, 07:59:06 ; Search time 151 Seconds  
(without alignments)  
3322.781 Million cell updates/sec

Title: US-09-943-075A-2

Perfect score: 3061  
Sequence: 1 MESLQGVFLVLLAAGLPQ.....PFSRGDRKDLLQDKPWL 572

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=vlh  
-O=/cgn2\_1/USPNO\_spool/US09943075/runat\_07042003\_155632\_18004/app-query.fasta\_1.711  
-DB=Published\_Applications\_NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09943075 -CGN\_1\_1\_57 -runat\_07042003\_155632\_18004  
-ICPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PTC\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PTCUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3061	100.0	2320	10 US-09-943-075A-1	Sequence 1, Appli
2	2700	88.2	2213	10 US-09-823-038A-27	Sequence 27, Appl
3	2695	88.0	1725	10 US-09-943-075A-7	Sequence 7, Appli
4	2163	70.7	2669	9 US-09-525-978B-82	Sequence 82, Appl

5	2162	70.6	2683	9 US-10-227-884-41	Sequence 41, Appl
6	2162	70.6	2683	9 US-10-230-163-41	Sequence 41, Appl
7	2162	70.6	2683	9 US-10-218-631-41	Sequence 41, Appl
8	2162	70.6	2683	9 US-10-230-338-41	Sequence 41, Appl
9	2162	70.6	2683	9 US-10-230-414-41	Sequence 41, Appl
10	2157	70.5	1683	10 US-09-943-075A-8	Sequence 8, Appli
11	2136	69.8	2661	9 US-10-039-272-1	Sequence 1, Appli
12	627.5	20.5	2130	9 US-09-812-2388-1	Sequence 1, Appli
13	627.5	20.5	2130	10 US-09-862-260A-1	Sequence 1, Appli
14	622	20.3	2131	9 US-10-047-539-1	Sequence 3, Appli
15	620	20.3	1881	9 US-10-047-539-1	Sequence 1, Appli
16	493	16.1	393	10 US-09-960-352-7359	Sequence 7359, Ap
17	462.3	15.1	318	10 US-09-864-761-18279	Sequence 18279, A
18	446.5	14.6	712	10 US-09-910-943-210	Sequence 210, App
19	405.5	13.2	478	10 US-09-864-761-1520	Sequence 1520, Ap
20	299	9.8	294	10 US-09-917-800A-139	Sequence 139, App
21	295.5	9.7	585	10 US-09-864-761-8827	Sequence 8827, Ap
22	295.5	9.7	585	10 US-09-864-761-8834	Sequence 8834, Ap
23	295	9.6	209	10 US-09-864-761-25253	Sequence 25253, A
24	293	9.6	209	10 US-09-864-761-25501	Sequence 25501, A
25	287	9.4	487	10 US-09-525-301-706	Sequence 706, App
26	257.5	8.4	474	10 US-09-960-352-5450	Sequence 5450, Ap
27	165	5.4	2567	9 US-09-992-598-240	Sequence 240, App
28	165	5.4	2567	9 US-09-989-293A-240	Sequence 240, App
29	165	5.4	2567	9 US-09-989-735-240	Sequence 240, App
30	165	5.4	2567	9 US-09-980-444-240	Sequence 240, App
31	165	5.4	2567	9 US-09-989-730-240	Sequence 240, App
32	165	5.4	2567	9 US-09-990-436-240	Sequence 240, App
33	165	5.4	2567	9 US-09-991-181-240	Sequence 240, App
34	165	5.4	2567	9 US-09-993-687-240	Sequence 240, App
35	165	5.4	2567	9 US-09-989-734-240	Sequence 240, App
36	165	5.4	2567	9 US-09-997-653-240	Sequence 240, App
37	165	5.4	2567	9 US-10-174-590-179	Sequence 179, App
38	165	5.4	2567	9 US-10-176-758-179	Sequence 179, App
39	165	5.4	2567	9 US-10-175-737-179	Sequence 179, App
40	165	5.4	2567	9 US-09-993-667-240	Sequence 240, App
41	165	5.4	2567	9 US-10-173-706-179	Sequence 179, App
42	165	5.4	2567	9 US-10-175-738-179	Sequence 179, App
43	165	5.4	2567	9 US-10-175-752-179	Sequence 179, App
44	165	5.4	2567	9 US-10-176-482-179	Sequence 179, App
45	165	5.4	2567	9 US-10-176-757-179	Sequence 179, App

#### ALIGNMENTS

RESULT 1  
US-09-943-075A-1  
; Sequence 1, Application US/09943075A  
; Patent No. US20020151486A1  
; GENERAL INFORMATION:  
; APPLICANT: Popoff, Steven N.  
; APPLICANT: Safado, Faye F.  
; APPLICANT: Owen, Thomas A.  
; APPLICANT: Smock, Steven L.  
; TITLE OF INVENTION: Osteoactivin Protein and Nucleic Acids Encoding the Same,  
; FILE OF INVENTION: Compositions and Methods of Stimulating Bone Differentiation  
; FILE REFERENCE: 71369,262  
; CURRENT APPLICATION NUMBER: US/09/943,075A  
; PRIOR FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: US 60/229,006  
; PRIOR FILING DATE: 2000-08-30  
; NUMBER OF SEQ IDS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2320  
; TYPE: DNA  
; ORGANISM: Rat osteoactivin  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (115)...(1833)  
US-09-943-075A-1

Alignment Scores:

Pred. No.: 7,63e-315 Length: 2320  
Score: 3061.00 Matches: 572  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-943-075a-2 (1-572) x US-09-943-075a-1 (1-2320)

QY 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaLaGlyLeuProLeuGln 20  
DB 115 ATGGAAGTCTCTGGGGGCTCTGGTATTCTGCTGTGCTGCAGGACTGCGGCTCCAG 174

QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
DB 175 GCGGCCACGGGTCCGGTGTGCTGGGCTTCCAGTATCCGATCATCATGAGGGAG 234

QY 41 AsnAsnGlnLeuArgGlyTyrPheSerSerAspGluAsnGlnTyrPheValPro 60  
DB 235 AACAAACCAATACGTGGCTGTCTTCAGATGAAATGAATGGGATGAACAGCTGTATCCA 294

QY 61 ValTyrArgArgGlyGluGlyArgTyrPheValValGlyValGlnAla 80  
DB 295 GTGTGAGGAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354

QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
DB 355 GCCCTAACCAAGTGTACCGGCTTGTGACCGGTATGCTTACAACTGGACGAGGGGAG 414

QY 101 ValPheProArgCysGlnGlyAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
DB 415 GTGTCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 474

QY 121 ArgSerAspLeuLeuAlaSerAspProTyrValTyrAsnTyrPheValValAlaAsp 140  
DB 475 AGAAGTGTATTTGGAGCTGGCTTGTGACCGGTATGCTTACAACTGGACGAGGGGAG 534

QY 141 AspGluAspTyrPheGluAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
DB 535 GATGAGGAGCTGGGAGAGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594

QY 161 ProPheProArgProHisGlyArgGlyLysTyrPheValTyrValPheHisThrLeu 180  
DB 595 CCCTTCCCTCGCCCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 654

QY 181 GlyGlnTyrPheGlnGlyLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
DB 655 GGTCAAGTATTTCAAAGCTGGGTCAGTGTTCAGCAGAGTTCCTATAAACACAGTCAAC 714

QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220  
DB 715 TTGACAGTTGGGCTTCAGGTCTATGGAAGTATGCTTTTCGAAGACACGCGCGGAGTAC 774

QY 221 IleProIleSerLysValValAspValTyrValIleThrAspGlnIleProIlePheVal 240  
DB 775 ATTCCCATCTCCAAAGTGAAGAGCTGTATGTGATAACAGATCAGATCCCTATATTTCGG 834

QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
DB 835 ACCATGTACCAAGAAATGACCGGAACTCGTCTGATGAAACCTTCTCAGAGACCTCCCC 894

QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
DB 895 ATTTCTTCGATGTCTCTATTCAGATCCCGGAGTTCCTTCCAACTACTCTCGCATTTCC 954

QY 281 TyrLysTyrPhePheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
DB 955 TACAAGTGGAACTTTGGGGAGCAACACTGGCTGTGTCTCTCCAACTATCAGATTTGAT 1014

QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
DB 1015 CACAGTATGTCTCAATGGAACTTCAACTTTAACTTCACTGCAAACTGCACTGCGG 1074

QY 321 GlyProCysProSerProThrProSerProSerSerSerSerSerSerProSerProAlaSer 340  
DB 1075 GGACCATGGCCCTCACCCACACTTTCGCTTCTTCTGACTTCTCCTTCGCTTCGCTCATCT 1134

QY 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyrLysSer 360  
DB 1135 TCGCCTTACCCACATATTAACACTAGTCCCTCTTAAATGCTTACTGGCTACAAATCC 1194

QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380  
DB 1195 ATGGAGCTGAGTGACATTTCCAATGAAACTGCCGAATAAAGATATGTTACTTCTCAGA 1254

QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400  
DB 1255 GCACCATCACAAATGTAGATGAATCTAGAAATCAACATCATCCAGGTAGCAGATGTC 1314

QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420  
DB 1315 CCAATCCCACTGCAGCTGCACAACTCACTGATGGACTTCAATGTGACCTGCAAAAGG 1374

QY 421 AlaThrProThrCluAlaCysThrIleSerAspProThrCysGlnIleAlaGlnAsn 440  
DB 1375 GCACCTCCCAAGAGCTGTAGATCATCTGACCCCACTGCCAGATGCCAGAAC 1434

QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460  
DB 1435 AGGTGTGAGCCCGGTGGCTGTGGATGAGCTGTGCCTCTGCTGAGGAGAGCCTTC 1494

QY 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480  
DB 1495 AATGGGTGCGGACGACTGTGTGAATTTCACTTGGGAGAGATGCAAGCTGGCCCTC 1554

QY 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500  
DB 1555 ACCAGCGCTGATCTCTATCCCTGGCAAGACCTAGCTCCCTCTGAGAACAGTGAAT 1614

QY 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
DB 1615 GGTGTCTGATCTTCAATGGCTGGCTGGCCATGTTGTGACCATGGTTACCATCTTGTCTG 1674

QY 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540  
DB 1675 TACAAAAACACAAGAGCTACAAAGCAATAGAAACTGCACCAAGACGTGGTCAAGGGC 1734

QY 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560  
DB 1735 AAAGGCTGAGTGTCTTCTCAGCCATGCAAAAGCCCGTTCTCCCGAGGAGACCGGGAG 1794

QY 561 LysAspProLeuLeuGlnAspLysProTyrPheMetLeu 572  
DB 1795 AAGGATCCACTGCTCCAGGACAAGCATGGATGCTC 1830

## RESULT 2

US-09-823-038A-27  
; Sequence 27, Application US/09823038A  
; Patent No. US2002005835A1  
; GENERAL INFORMATION:  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Slesman, Matthew  
; APPLICANT: Abernethy, Nevin  
; APPLICANT: Onrust, Rene  
; APPLICANT: Kumble, Anand  
; APPLICANT: Murison, Greg  
; TITLE OF INVENTION: Compositions Isolated From Stromal Cells  
; FILE REFERENCE: 11000.1037c3  
; CURRENT APPLICATION NUMBER: US/09/823,038A  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 2213  
; TYPE: DNA



ORGANISM: Mouse  
US-09-823-038A-27

Alignment Scores:

Pred. No.: 1,51e-276 Length: 2213  
Score: 2700.00 Matches: 507  
Percent Similarity: 92.29% Conservative: 20  
Best Local Similarity: 88.79% Mismatches: 42  
Query Match: 88.21% Indels: 2  
DB: 10 Gaps: 1

US-09-943-075A-2 (1-572) x US-09-823-038A-27 (1-2213)

QY 1 MetGluSerLeuGlyValLeuValPheLeuLeuLeuAlaAlaClyLeuProLeuGln 20  
DB 23 ATGGAAGTCTCTCGGGGGCTGGGATTTCTGCTGCTGGCTGCAGGACTGCTCTCCAG 82  
QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
DB 83 CTTGCCAAGCGATTTCTGTGATGCTGGGCCATGAACATGATCTCCCAATCATATGAGAG 142  
QY 41 AsnAsnGlnLeuArgGlyTyrPsrSerAspGluAsnGlnTyrAspGluGlnLeuTyrPro 60  
DB 143 CACAACCAATTTACCTGGCTGCTGCGATGAAATGAATGAATGAATGAATGAATGAATGA 202  
QY 61 ValTyrArgGlyGluGlyArgTyrLysAspSerTrpGluGlyGlyArgValGlnAla 80  
DB 203 GTTGGAGAGGGAGGACGACGAGTGGAGGACTCTCTGGGAAGGAGGCGGTGTGCAGGCA 262  
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
DB 263 GTCTGACCACTGACTCACCGGCTCTGGTGGGTTCCTCAATATACCTTTTGGTGAACCTG 322  
QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
DB 323 GTTGTTCCTCCAGATCCAGAGAGGAGATGCTAATGGCAATATCTGTATGAGAGAACTGC 382  
QY 121 ArgSerAspLeuLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140  
DB 383 AGGAATGATTTGGGACTGACTCTGACCTGCATGCTCACTGACCTGACGAGGCGACAT 442  
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160  
DB 443 GATGTGACTGGGAAGATGGCAGCCAGGAGCCAGCATCTCAGGTTCCTCCGACAGGAG 502  
QY 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
DB 503 CCTTCCCTCGCCCCCATGGATGGAGAAATGGAGCTTTGTCTACGCTCTTCACACACTT 562  
QY 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
DB 563 GGCAGATATTTCCAAAACCTGGGTGCTGACGACGGGTTCATATAACACACAGTCAAC 622  
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220  
DB 623 TTGACAGCTGGCCCTCAGGTGATGGAAGTGTCTTTCGAAGATACGGCGGCGCATAC 682  
QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
DB 683 ATTCCCATCTCGAAGGTGAAGATGTGTATGTATTAACAGATCAGATCCCTGTATTCGTG 742  
QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
DB 743 ACCATGTCACAGAAATGACAGAACTGCTGTGATGAGATCTTCTCAGAGACTCTCC 802  
QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
DB 803 ATCTGCTTCGATGCTCATTCATGATCCAGGACCTTCTCAACAGCACTCTGCCATTTCC 862  
QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn 300  
DB 863 TACAAGTGAACCTTTGGGGACACACTGGGCTGTTGTCTCTCCAAACATCACACTTTGAAT 922

QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
DB 923 CACACTTATGTCTCAATGGAACTTCAACCTTAACCTCACTCACTCACTCACTCACTCACT 982  
QY 321 GlyProCysProSerProThrProSer-----ProSerSerSerThrSerProSerPro 338  
DB 983 GGGCCATGCCCTGCCCTTCCGCTTCGACTCCGCTCCGCTTCAACTCCGCTCCGCTCCGCT 1042  
QY 339 AlaSerSerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 358  
DB 1043 CCGCCCTCACCTCTGCCACCATTTATCAACACCTAGCCCTCTTTAATGCTACTGTTAC 1102  
QY 359 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 378  
DB 1103 AAATCATGGAGCTGAGTGACATTTCCAATGAAACTGCGGAATAAACAAGATGCTAC 1162  
QY 379 PheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAla 398  
DB 1163 TTCAGAGCCACCATCAATTTGTAGAGGGATCTCGAAGTCAAGATCAATGAGATGAGCA 1222  
QY 399 AspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCys 418  
DB 1223 GATGTCCCATGCCACACCGCAGCTGCCAACTCCCTGATGGACTTCACTGTGACCTGC 1282  
QY 419 LysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 438  
DB 1283 AAGSGGCGCACCCCATGGAAGCTGTACGATCATCTCCGACCCACCTGCCAGATCGCC 1342  
QY 439 GlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArg 458  
DB 1343 CAGAACCGGCTGTCAGCGCTGTGCTGTGGATGGCTGTGCTGTGCTGTGCTGTGAGAAGA 1402  
QY 459 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 478  
DB 1403 GCTTCAATGGGTCTGGCAGCTCTGCTGTGAATTTCACTCTGGAGATGATGCAAGCCTG 1462  
QY 479 AlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThr 498  
DB 1463 GCGCTCACCGACCGCTGATCTATCCCTGGCAAGACCCAGACTCCCTCTGAGAGCA 1522  
QY 499 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIle 518  
DB 1523 GTGAATGGTCTCTGATCTCCATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1582  
QY 519 LeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValVal 538  
DB 1583 TTGCTGTACAAAAACACCAAGCGCTACAAAGCAATAGGAACTGCCCCAGGAACACGGTC 1642  
QY 539 LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAsp 558  
DB 1643 AAGGGCAAGGCGCTGAGTGTCTCTCAGCCACGCGAAGCCCGTCTTCTCCGAGGAGAC 1702  
QY 559 ArgGlyLysAspProLeuLeuGlnAspLysPro 569  
DB 1703 CAGGAGAAGGATCCATTCTCCAGGACAAAGCA 1735

RESULT 3

US-09-943-075A-7

; Sequence 7, Application US/09943075A  
; Patent No. US20020151486A1  
; GENERAL INFORMATION:  
; APPLICANT: Popoff, Steven N.  
; APPLICANT: Safado, Faye F.  
; APPLICANT: Owen, Thomas A.  
; APPLICANT: Smock, Steven L.  
; TITLE OF INVENTION: Osteoactivin Protein and Nucleic Acids Encoding the Same,  
; TITLE OF INVENTION: Compositions and Methods of Stimulating Bone Differentiation  
; FILE REFERENCE: 71369.262  
; CURRENT APPLICATION NUMBER: US/09/943.075A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: US 60/229,006  
; PRIOR FILING DATE: 2000-08-30  
; NUMBER OF SEQ ID NOS: 8





APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Smith, Victoria  
APPLICANT: Stephan, Jean-Philippe F.  
APPLICANT: Watanabe, Colin L.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3530P1C79  
CURRENT APPLICATION NUMBER: US/10/227,884  
CURRENT FILING DATE: 2002-08-26  
PRIOR APPLICATION NUMBER: 10/119,480  
PRIOR FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063549  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/064103  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/069873  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086392  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089905  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090691  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/095302  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/095318  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/095916  
PRIOR FILING DATE: 1998-08-10  
PRIOR APPLICATION NUMBER: 60/096146  
PRIOR FILING DATE: 1998-08-11  
PRIOR APPLICATION NUMBER: 60/096791  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 60/097986  
PRIOR FILING DATE: 1998-08-26  
PRIOR APPLICATION NUMBER: 60/098544  
PRIOR FILING DATE: 1998-08-31  
PRIOR APPLICATION NUMBER: 60/099596  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 60/099598  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 60/099803  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099811  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099812  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099816  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/100038  
PRIOR FILING DATE: 1998-09-11  
PRIOR APPLICATION NUMBER: 60/100385  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100390  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100627  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100848  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/100919  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/101477  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101738  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101741  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101786  
PRIOR FILING DATE: 1998-09-25  
PRIOR APPLICATION NUMBER: 60/101916  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101922  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/106178  
PRIOR FILING DATE: 1998-10-28  
PRIOR APPLICATION NUMBER: 60/106248  
PRIOR FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: 60/106464  
PRIOR FILING DATE: 1998-10-30  
PRIOR APPLICATION NUMBER: 60/106905  
PRIOR FILING DATE: 1998-11-03  
PRIOR APPLICATION NUMBER: 60/108787  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: 60/108801  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: 60/108849  
PRIOR FILING DATE: 1998-11-18  
PRIOR APPLICATION NUMBER: 60/112422  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/113296  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/113605  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113621  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/115558  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115565  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115733  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119549  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/123618  
PRIOR FILING DATE: 1999-03-10  
PRIOR APPLICATION NUMBER: 60/125259

;  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 60/125775  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: 60/126773  
; PRIOR FILING DATE: 1999-03-29  
; PRIOR APPLICATION NUMBER: 60/127887  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 60/130232  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR APPLICATION NUMBER: 60/131022  
; PRIOR FILING DATE: 1999-04-26  
; PRIOR APPLICATION NUMBER: 60/131270  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131291  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131445  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: 60/134287  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/140650  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/140723  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/141037  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 60/144758  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: 60/146222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: 60/146963  
; PRIOR FILING DATE: 1999-08-03  
; PRIOR APPLICATION NUMBER: 60/149320  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/149638  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/151733  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/164418  
; PRIOR FILING DATE: 1999-11-09  
; PRIOR APPLICATION NUMBER: 60/166361  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 60/169445  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/169495  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/169835

## Alignment Scores:

Pred. No.:	2,71e-219	Length:	2683
Score:	2162.00	Matches:	396
Percent Similarity:	82.72%	Conservative:	78
Best Local Similarity:	69.11%	Mismatches:	87
Query Match:	70.63%	Indels:	12
DB:	9	Gaps:	2

US-09-943-075a-2 (1-572) x US-10-227-884-41 (1-2683)

Qy	1	MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaAaGlyLeuProLeuGln	20
Db	82	ATGGAATGCTCTACTATTCTCGGGATTCTGCTCGCTGGCTGCAAGATTGCCACTTGAT	141
Qy	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu	40
Db	142	CCCCCAACAGTTTCATGATGTCGGCAATGAAGACCTTCCTGCTTCATGAGGGAG	201
Qy	41	AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro	60
Db	202	CACAATCAATAAATGGCTGGCTTCATGATGAAATGACTGGAATGAAAACTACCCA	261
Qy	61	ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla	80

Db	262	GTJTGGAAGCGGGAGACATGAGGTGGAAGAACTCCTGGAAGGAGCGCGCTGTGACGGCG	321
Qy	81	AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
Db	322	GTCTGACCAAGTGTCTACCAAGCCCTCGTGGCTCAATATACATTTTGGCGGTGAACCTG	381
Qy	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys	120
Db	382	ATATTCCCTAGATGCCAAAGGAAGATGCCAATGCCAATAGCTATGATGAGAAAGAACTGC	441
Qy	121	ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrThrGlyAlaAsp	140
Db	442	AGAAATGAGGCTGGTTATCTCTCTCATCGTATGTACAACTGGACAGCATGGTCAGAG	501
Qy	141	AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys	160
Db	502	GACAGTGACGGGGAAATGGCACCGGCAAGAACCCATACATACGCTTCCCTGATGGAA	561
Qy	161	ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu	180
Db	562	CCITTTCTCACCCCGGATGGAGAAGATGGAATTTTCATCTACGCTCTTCCACACACTT	621
Qy	181	GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn	200
Db	622	GGTCAGTATTTCCAGAAATGGACGATGTCAGTGAGAGTTTCTGTGAACACAGCAAT	681
Qy	201	LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr	220
Db	682	GTGACACTTGGCGCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGCATAT	741
Qy	221	IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal	240
Db	742	GTTTCCATCGCACACAGTGAAGATGTGTACGTGGTAAACAGATCATCTCTGTGTGTG	801
Qy	241	ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro	260
Db	802	ACTATGTTCCAGAAAGACATCGAATTCACGACGAACCTTCCCAAGATCTCCCC	861
Qy	261	IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer	280
Db	862	ATTATGTTTGTATGTCCTGATTCATGATCTAGCCACTTCCCTCAATTTATCTACCAAT	921
Qy	281	TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn	300
Db	922	TACAGTGGAGCTTCGGGATTAATACTGCCCTGTTGTTCCACCAATCATCTGTGAAT	981
Qy	301	HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro	320
Db	982	CACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCCTGTGAAAGCTGCAGCACCA	1041
Qy	321	GlyProCysProSerProThrProSerProSerSerSerSerSerProSerProAlaSer	340
Db	1042	GGACCTTGTCCGCCACCGCCACCCACCCACCCAGCA-----CCTTCA	1080
Qy	341	SerProSerProThrLeuSerThr-----ProSerProSerLeuMetPro	355
Db	1081	AAACCCACCCCTTCTTTAGCACTACTTAANTCTTATGATCAACACCCACCCAGGACCT	1140
Qy	356	ThrGlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArg	375
Db	1141	ACTGTCACAAACCCCTGGAGCTGAGTAGGATCTCTGATGAAAGCTGCCAGATTACAGA	1200
Qy	376	TyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle	395
Db	1201	TATGSCCACTTTCACGCCACCATCACATTTGTAGAGGAATCTTAGAGGTTAACATCATC	1260
Qy	396	GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle	415
Db	1261	CAGATGACAGACGCTCTGATGCGCGTGGCATGGCTGAAAGCTCCCTAATAGACTTTGTC	1320
Qy	416	ValThrCysLysGlyAlaThrProThrGluAlaLacCysThrIleIleSerAspProThrCys	435
Db	1321	GTGACGTGCCAAGGGAGCATTCACAGGGAGGTCTGTACCATCATTTCTGACCCCACTGC	1380

```
Oy 436 GhlleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSer 455
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
Db 1381 GAGATCACCCAGAACAGTCTGCACGCTCTGGATGTGGATGGATGTGTCTGCTGACT 1440
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
Oy 456 ValArgArgAlaPheAsnGlySerGlyThrTyrcysValAsnPheThrLeuGlyAspAsp 475
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
Db 1441 GTGAGACGACCTTCATGTGGTCTGGGACGTACTGTGTGAACCTCACCTCGGGGGATGAC 1500
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
Oy 476 AlaSerLeuAlaThrSerAlaLeuLeuSerIleProGlyLysAspLeuGlySerPro 495
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
Db 1501 ACAAGCTGGCTCTCAGACGACCCCTGATTTCTCTTCAGACAGACCCAGCCTCGCCT 1560
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089532
Oy 496 LeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMet 515
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
Db 1561 TTAAGGATGGCAACAGTGCCTGATCTCCGTTGGTGTGGCCATATTGTCACTGTG 1620
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089905
Oy 516 ValThrIleLeuLeuTyrcysLysLysThrTyrcysLysProIleGlyAsnCysThrArg 535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
Db 1621 ATCTCCCTCTGGTGTACAAAACACAAAGGAATACACCCCAATAGAAAATAGTCTCTGGG 1680
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090557
Oy 536 AsnValValLysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSer 555
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090691
Db 1681 AATGTGTGTGACAGCAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTTC 1740
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
Oy 556 ArgGlyAspArgGlyLysAspProLeuLeuGlnAspLys 568
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
Db 1741 CCGGAACACCGAAGGAAGGATCCGCTACTCAAAAACCCAA 1779
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791

RESULT 6
US-10-230-163-41
; Sequence 41, Application US/10230163
; Publication No. US20030036635A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P5530FIC96
; CURRENT APPLICATION NUMBER: US/10/230,163
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
```

[illegible]





Db 382 ATATTCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAACTGC 441  
Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAsp 140  
Db 442 AGAATAGAGGTGGTTTATCTGCTGATCGGTATGTTTACAAGTGGACAGCATGTCAGAG 501  
Qy 141 AspGluAspTrpGluAspAsnTrpSerGlnGlnHisLeuArgPheProAspGlyLys 160  
Db 502 CACAGTGAAGGGGAAATGGCACCAGCCCAAGCCCATCATAGCTTCTCCCTGATGGGAA 561  
Qy 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
Db 562 CTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTAGCTCTCCACAGACTT 621  
Qy 181 GlyGlnTyrPheGlnLysLeuGlnLysLeuGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
Db 622 GGTCAATATTTCCAGAAATTTGGACGATGTTCAAGTCAAGATTTCTGTCAACACAGCCCAT 681  
Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220  
Db 682 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTCTCTACAGAAAGACATGGACGGCATAT 741  
Qy 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
Db 742 GTTCCATCGCACAAAGTGAAGATGCTACGTGGTAACAGATCAGATCCTGTGTGTG 801  
Qy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
Db 802 ACTATGTTCAGAGAAACGATCGAAATTCATCGGAGAAACCTTCTCAAAAGATCTCCCC 861  
Qy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
Db 862 ATTATGTTTGTATGCTGATTCATGATCTAGTCCCTAGCCACTTCTCAATATTCTACCAATTAA 921  
Qy 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValIleSerAsnHisThrLeuAsn 300  
Db 922 TACAAGTGGAGCTTCGGGGATAATCTAGTGGCTGTGTGTTCCACCAATCATCTAGTGAAT 981  
Qy 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
Db 982 CACAGTATGCTCAATGGAACCTTCAGCCTTAACCTCAGTGTGAAGCTGCAGCACCA 1041  
Qy 321 GlyProCysProSerProThrProSerProSerProSerProSerProSerProSerPro 340  
Db 1042 GGACCTTGTCCGCCACCCGCCACCCACCCAGCA-----CCTTCA 1080  
Qy 341 SerProSerProThrLeuSerThr-----ProSerProSerLeuMetPro 355  
Db 1081 AAACCCACCCCTCTTTAGCACTACTCTAAATCTTATGATTAACACCCAGGACCT 1140  
Qy 356 ThrGlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArg 375  
Db 1141 ACTGGTGACACCCCTGGAGCTGAGTAGGATTCCTGATGAACCTGCAGATTAACAGA 1200  
Qy 376 TyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle 395  
Db 1201 TATGCCACTTCAAGCCACCATCACAATTTGTAGAGGGAATCTTAGAGGTTAAACATCATC 1260  
Qy 396 GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415  
Db 1261 CAGATGACAGACGCTGCTGCTGCCGTCCTGATGCTGATGCTGATGCTGATGCTGATGCT 1320  
Qy 416 ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435  
Db 1321 GTGACCTGCCAAGGAGCATTCGCCAGGAGGCTGTGTACCATCATTTCTGACCCCATGTC 1380  
Qy 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuLeuSer 455  
Db 1381 GAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGTGGATGTGGATGTGGATGT 1440  
Qy 456 ValArgArgAlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAsp 475  
Db 1441 GTAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTTGGGGATGAC 1500  
Qy 476 AlaSerLeuAlaLeuThrSerAlaLeuLeuSerIleProGlyLysAspLeuGlySerPro 495  
Db 1501 ACAAGCTGGCTCTCACGAGCACCTGATTTCTGTTCCTGACAGAGACCCAGCCTCGCCT 1560  
Qy 496 LeuArgThrValAsnGlyValLeuLeuSerIleGlyCysLeuAlaMetPheValThrMet 515  
Db 1561 TTAAGGATGGCAACAGTGCCTGATCTCGTGGCTGCTTGGCCATATTTGTCACTGTG 1620  
Qy 516 ValThrIleLeuLeuTyrLysHisLysThrTyrLysProIleGlyLysCysThrArg 535  
Db 1621 ATCTCCCTCTTGGTGTACAAAAACAAGGAATACACCACTAGAAATAGTCTCTGGG 1680  
Qy 536 AsnValValLysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaPropheSer 555  
Db 1681 AATGTGTGCAGAACCAAGCCCTGAGTGTCTTCTCAACCGTGCAGAAAGCGGTGTCTTC 1740  
Qy 556 ArgGlyAspArgGluLysAspProLeuLeuGlnAspLys 568  
Db 1741 CCGGAAACCAAGGAAGGATCCGCTACTCAAAACCAA 1779

## RESULT 8

US-10-230-338-41  
; Sequence 41, Application US/10230338  
; Publication No. US20030044934A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3530P1C92  
; CURRENT FILING DATE: 2002-08-28  
; PRIOR APPLICATION NUMBER: US/10/230,338  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 246  
; SEQ ID NO 41  
; LENGTH: 2683  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-230-338-41

Alignment Scores: 2.71e-219 Length: 2683  
Pred. No.:

Score:	2162.00	Matches:	396
Percent Similarity:	82.72%	Conservative:	78
Best Local Similarity:	69.11%	Mismatches:	87
Query Match:	70.63%	Indels:	12
DB:	9	Gaps:	2

US-09-943-075A-2 (1-572) x US-10-230-338-41 (1-2683)

Qy	1	MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln	20
Db	82	ATGGAAATGCTCTACTATTTCTGGGATTTCTGCTCTGGCTGCAAGATGGCACTGAT	141
Qy	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu	40
Db	142	CGCGCCAAACGATTTCATGATGCTGGCAATGAAGACCTTCTGCTTACATGAGGAG	201
Qy	41	AsnAsnGlnLeuArgGlyTyrPsrSerAspGluAsnGluTrpAspGluGlnLeuTyrPro	60
Db	202	CACAATCAATTAATAGGCTGGCTCTCTGATGAAATGACTGGAATGAAAACTCACCCA	261
Qy	61	ValTTPArgArgGlyGluGlyArgTgPLysAspSerTrpGluGlyGlyArgValGlnAla	80
Db	262	GTGTGGAGCGGGAGACATGAGTGGRAAACTCTGGAAGGAGCGCGTGTGCAGCGC	321
Qy	81	AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
Db	322	GTCTGTGACAGTGACTCACCAGCCCTCTGGGCTCAATATAACATTTGCGGTGAACCTG	381
Qy	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys	120
Db	382	ATATTCCCTACATGCCAAAGAGAGATGCCAATGCCACATAGTCTATGAGAAGAACTGC	441
Qy	121	ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp	140
Db	442	AGAAATGAGCGTGGTTATCTGCTGATCCGTTATGTTTACAACATGGACATGGTCAGAG	501
Qy	141	AspGluAspTTPGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys	160
Db	502	GACAGTGGCGGGAATATGGCACCAGCCAAACCCATCAATACGTCTCTCCATGGGAAA	561
Qy	161	ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu	180
Db	562	CCTTTTCTCCACCACCCCGGATGGAGAATGGAAATTTCACTACGTCTCCACACACTT	621
Qy	181	GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn	200
Db	622	GGTCAGTATTTCCAGAAATTTGGACAGCTGTTCACTGAGAGTTCTGTGAACACACCCAAT	681
Qy	201	LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr	220
Db	682	GTCAACACTTGGCGCTCAACTCATGGAAAGTGACTGTCTACAGAAGACATGGAGCGGCATAT	741
Qy	221	IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal	240
Db	742	GTTCCTCCATCGCACAAAGTGAAGATGTGTACGTGGTACAGATCAGATTTCTGTGTGTTGTG	801
Qy	241	ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro	260
Db	802	ACTATGTTCACAGAAGACGATCGAAATTCATCCGACGAAACCTTCCTCAAAAGATCTCCCC	861
Qy	261	IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer	280
Db	862	ATTATGTTTGATGCTTGATTCATGATCTCCTAGCCACTTCCTCAATATTCTTACCATTATAC	921
Qy	281	TyrLysTyrAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn	300
Db	922	TACAAGTGGAGTTCGGGGATAATACTGGCCGTGTTGTTTCCACCAATCATACTCTGAAT	981
Qy	301	HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuLeuThrValGlnThrAlaValPro	320
Db	982	CACAGCTATGTGCTCAATGGAACTTCAGCCCTTAACCTCACTGTGAAAGCTGCACGACA	1041
Qy	321	GlyProCysProSerProThrProSerProSerSerSerSerSerSerSerProAlaSer	340

Db	1042	GGACCTTGTCCGCAACCGCCACCAACCCAGA-----CCTTCA	1080
Qy	341	SerProSerProThrLeuSerThr-----ProSerProSerLeuMetPro	355
Db	1081	AAACCCACCCCTTCTTTAGCAACTACTCTAAAATCTTAGATTCAACACACCCAGGACCT	1140
Qy	356	ThrGlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArg	375
Db	1141	ACTGGTGCACACCCCTGGAGCTGAGTAGAGTCTTGATGAACAACTGCCAGATTAAACAGA	1200
Qy	376	TyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle	395
Db	1201	TATGGCCACTTTTCAAGCCACCATCAATTTGTAGAGGAATCTTAGAGGTTTAACATCATC	1260
Qy	396	GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle	415
Db	1261	CAGATGACAGAGCTCTGATGCCGGTCCCATGGCTGAAGAGTCCCTAATAGACTTTGTCT	1320
Qy	416	ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleLeuSerAspProThrCys	435
Db	1321	GTGACCTGCCAAGGAGGACATTCGCCAGGAGTCTGTACCATCATTTCTGACCCCACTGC	1380
Qy	436	GlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSer	455
Db	1381	GAGATCACCCAGACACACAGCTCCACCCCTGTGGATGTGGATGTGTCTGCTGACT	1440
Qy	456	ValArgArgAlaPheAsnGlySerGlyThrCysValAsnPheThrLeuGlyAspAsp	475
Db	1441	GTGAGACCAACCTTCAATGGGTCTGGAGCTACTGTGTGAACCTCACCCCTGGGGATGAC	1500
Qy	476	AlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerPro	495
Db	1501	ACNAGCCTGGCTCTCAGACGACCCCTTGATTTCTGTTCTTGACAGAGACCCGCTCGCCT	1560
Qy	496	LeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMet	515
Db	1561	TTAAGGATGGCAACACGTGCCCTGATCTCCGTGGCTGTGGCCATATTTGTCTACTGTG	1620
Qy	516	ValThrIleLeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArg	535
Db	1621	ATCTCCCTCTGGGTACAAAAACCAAGGAATACACCCCAATAGAAAATAGTCTCTGGG	1680
Qy	536	AsnValValLysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSer	555
Db	1681	ATGTGGTGCAGAACCAAGGCCCTGAGTGTCTTCTCAACCGTGCAAAAGCGGTGTTCTTC	1740
Qy	556	ArgGlyAspArgGluLysAspProLeuLeuGlnAsnLys	568
Db	1741	CGGGGAAACCAAGGAAAGGATCGCTACTCAAAAACCAA	1779

RESULT 9  
US-10-230-414-41  
? Sequence 41, Application US/10230414  
? Publication No. US20030050448A1  
? GENERAL INFORMATION:  
? APPLICANT: Baker, Kevin P.  
? APPLICANT: Desnoyers, Luc  
? APPLICANT: Gerritsen, Mary  
? APPLICANT: Goddard, Audrey  
? APPLICANT: Godowski, Paul J.  
? APPLICANT: Grimaldi, J. Christopher  
? APPLICANT: Gurney, Austin L.  
? APPLICANT: Smith, Victoria  
? APPLICANT: Stephan, Jean-Philippe F.  
? APPLICANT: Wood, William L.  
? APPLICANT: Watanabe, Colin L.  
? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
? FILE REFERENCE: P3530P1C98  
? CURRENT APPLICATION NUMBER: US/10/230,414  
? CURRENT FILING DATE: 2002-08-28  
? PRIOR APPLICATION NUMBER: 10/119,480

Oy	181	GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn	200
Db	622	GGTCAGTAATTCCAGAAATTTGGGACGAGTGTTCTAGTCGAGAGTTTCTGTGTGAACACGAGCAAT	681
Oy	201	LeuThrValGlyProGlnValMetGluValIleValPheArgGHisGlyValArgAlaTyr	220
Db	682	GTACACACTTGGGCCTCAACTCATGGAAGTAGCTGTCTACAGNAGACATGGACGGGCATAT	741
Oy	221	IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal	240
Db	742	GTTCOCATCGCACAGTGAAGAAGATGTGTACGTGGTAACAGATCGAGATTCTCTGTGTGTG	801
Oy	241	ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro	260
Db	802	ACCATGTGTCGAGAAGCAGATCGNAATTATCCGAGGAACCTTCCTCAAGATCTCCCC	861
Oy	261	IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer	280
Db	862	ATTATGTTTGTATGTCCTGATTTCATGATCTTCAGCCACTTCCTCAATATTCTTACCATAAAT	921
Oy	281	TyrLysTrpAsnPhelGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn	300
Db	922	TACAAGTGGAGCTTCGGGATTAATACTGGCGCTGTTTGTTCACCAATCATACTGTGAAT	981
Oy	301	HisThrTyrValLeuAsnGlyThrPheAsnPhaAsnLeuThrValGlnThrAlaValPro	320
Db	982	CACACGTATGTCTCAATGGAACTTCAGCCTTTAACCTCACTGTGAAGCTGCAGCACCA	1041
Oy	321	GlyProCysProSerProThrProSerProSerSerSerSerSerProAlaSer	340
Db	1042	GGACCTTGTCGCCGCCGCCACCACCACCCAG-----CCTTCA	1081
Oy	341	SerProSerProThrLeuSerThr-----ProSerProSerLeuMetPro	355
Db	1081	AARCCACCCCCTCTTTAGCAACTACTCTAAAATCTTATGATTCAAACCCCGAGGACCT	1141
Oy	356	ThrSlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArg	375
Db	1141	ACTGSGTGACACCCCTCGAGCTCAGTAGGATTCCTTGATGAAACCTGCAGATTAAACAG	1201
Oy	376	TyrGlyTyrPheArgNaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle	395
Db	1201	TATSGGCACCTTTCAGCCACCCTCCACAAATGTAGAGGGAACTTAGAGGTTTACATCATC	1261
Oy	396	GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle	415
Db	1261	CAGATCAGACAGAGCTCCTGATCGCCGTGCCATGGCCTTGAAGCTCCCTAANTAGACTTTGT	1321
Oy	416	ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys	435
Db	1321	GTGACCTGCCAAGGAGGATTCGCCAGGAGGCTGTGTACCATCATTTCTGACCCCACTGC	1381
Oy	436	GlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSer	455
Db	1381	GAGATCACCCAGACACAGTCTGCAGCCCTGTGGATGTGGATGTAGATGTGCTGCTGACT	1441
Oy	456	ValArgAlaPheAsnGlySerGlyThrTyrCysValAsnPhaThrLeuGlyAspAsp	475
Db	1441	GTGAGACGAACCTTCAATGGGTCTGGGACGCTACTGTGTGAACCTCACCTGGGGGATGAC	1501
Oy	476	AlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerPro	495
Db	1501	ACAAGCCTGGCTCTCACGAGACCCGTGATTTCTGTTCCTGCAGAGACCCAGCCTCGCCT	1561
Oy	496	LeuArgThrValIasnGlyValIleuIleSerIleglyCysLeuAlaMetPheValThrMet	515
Db	1561	TTAAGCATGGCAACAGTCAGTCTGATCTCGTGGCTGTCTGGCCATATTTGTCACTGTG	1621
Oy	516	ValThrIleLeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArg	535
Db	1621	ATCTCCCTCTGGTGTACAAAAACACAGGAATACACCCCAATAGAAATAGTCTGGG	1681
Oy	536	AsnValValLysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaIapropheSer	555

[illegible]

[illegible]

	Qy	321	GlyProCysProSerProThrProSerProSerProSerThrSerProSerProAlaSer	340
	Dy	961	GGACCTTTCGCCACCAGCCACCACCACCGACCT	996
	Qy	341	SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyrrLysSer	360
	Dy	997	-----TCMAAACCCACCCCTCTTAGGNACCTGTGGTGACACCCC	103
	Qy	361	MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgrTyrrGlytyrPheArg	380
	Dy	1039	CTGGAGCTGAGTAGAGTTCTGTATGAATAACTGCCAGATTAAACAGATAGGCTACTTTC	109
	Qy	381	AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAaspVal	400
	Dy	1099	GCCACCATCACANTGTAGAGGANCTTAGAGGTTTAACATCATCCAGACGACGC	115
	Qy	401	ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly	420
	Dy	1159	CTGATCGCGTGCCTGAGCCTGAAAGCTCCTCAATAGACTTTGTGTCGACCTGCCAAGG	121
	Qy	421	AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn	440
	Dy	1219	AGCATTCACACGGAGGCTGTGACCATCATTTCTGACCCACCCTCGAGATCACCCAGA	127
	Qy	441	ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArgAlaPhe	460
	Dy	1279	ACAGTCTCGACCCCTGGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAAC	133
	Qy	461	AsnGlySerGlyThrTyrrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu	480
	Dy	1339	ATGGGTCTGGAGAGTACTGTGTGAACCTTACCCTGGGGATGACACNAGCCTGGCTCT	139
	Qy	481	ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgrThrValAsn	500
	Dy	1399	ACGACACCCCTGATTCTGTTCCTGCAGACAGACCCAGCCTCGCCTTTAAGGATGGCAA	145
	Qy	501	GlyValLeuIleSerIleGlyLeuAlaMetPheValThrMetValThrIleLeuLeu	520
	Dy	1459	AGTGCCCTGATCTCCGTTGGCTGCTGGCCATATTGTGCATGTGATCTCCCTTGGTG	151
	Qy	521	TyrLysLysHisLysThrTyrrLysProIleGlyAsnCysThrArgAsnValAlaLysGly	540

361	MetGluLeuSerAspIleSerAsnGlnAsnGlyGlyGlyPheArg	1039	CtGGAGCTGAGTAGATTCTCGATGAAATCTGCAGATTAAACAGTAAGGCTACTTTCAA	109
Qy		Db		
381	AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal	1099	GCCACCATCAAAATTGTAGAGGAATCTTAGAGGTAAACATCATCCAGATGACAGACGTC	115
Qy		Db		
401	ProIleProThrLeuGlnProAsnSerLeuMetAspPheIleValThrCysLysGly	1159	CTGATGCCGGTCCCATGGCTTGAAGAGCTCCATATAGACTTTGTCTGTGACCTGCCAAGGG	121
Qy		Db		
421	AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn	1219	AGCATTCGCCAGGAGGTCTGTACCATCATTTCTGCACCCACCTCGAGATCACCCGAAC	127
Qy		Db		
441	ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArgAlaPhe	1279	ACAGTCTGCACCCCTGTGGATGGATGAGATGTCTCTGCTGACTGTGAGACGAACCTTC	133
Qy		Db		
461	AsnGlySerGlyThrThrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu	1339	AATGGTCTGGAGAGTACTGTGTGAACCTCACCTCGGGGATGACACAAAGCTGGCTCTC	139
Qy		Db		
481	ThrSerAlaLeuIleSerIleProGlyIleAspLeuGlySerProLeuLeuArgThrValAsn	1399	ACGACGACCCGTATTCTGTCTCTGCAGACGACCCAGCCCTCGCCTTTAAGGATGGCAAC	145
Qy		Db		
501	GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu	1459	AGTGCCCTGATCTCCGTGGCTGCTTGGGCATATTGTTCACGTGATCTCCCTCTGGTG	151
Qy		Db		
521	TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly			

1519 TACAAAAACACAGGAATACACCAATAGAAATAGCTCGGAATGTGTCAGAACG 1578  
 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560  
 1579 AANGCCCTGAGTGTCTTCTCAACCGTGCAGAAAGCGTGTCTTCCCGGGAACAGGAA 1638  
 561 LysAspProLeuLeuGlnAspLys 568  
 1639 AAGGATCCGCTACTCAAAAACCAA 1662

RESULT 11

US-10-039-272-1  
 ; Sequence 1, Application US/10039272  
 ; Patent No. US20020168653A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RAMESHWAR, Pranela  
 ; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE  
 ; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey  
 ; CURRENT APPLICATION NUMBER: US/10/039,272  
 ; CURRENT FILING DATE: 2001-10-20  
 ; PRIOR APPLICATION NUMBER: US 60/241,881  
 ; PRIOR FILING DATE: 2000-10-20  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 2661  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-039-272-1

Alignment Scores:  
 Pred. No.: 1,54e-216 Length: 2661  
 Score: 2136.00 Matches: 293  
 Percent Similarity: 82.39% Conservatives: 75  
 Best Local Similarity: 69.19% Mismatches: 15  
 Query Match: 69.78% Indels: 15  
 DB: 9 Gaps: 1

US-09-943-075A-2 (1-572) x US-10-039-272-1 (1-2661)

QY 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20  
 DB 60 ATGGAAATGCTCTACTATTTCTCGGGATTTCTGCTCTGCTGCAAGATGCACTTGTAT 119  
 QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
 DB 120 GCGGCAACACGATTTCAATGATGCTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179  
 QY 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro 60  
 DB 180 CACATCAATTAATGGCTGCTCTCTGATGAAATGACTGGAAATGAAACTCTACCCA 239  
 QY 61 ValTrpArgGlyGlyGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla 80  
 DB 240 GTGTGGAACGGGGAGACATGATGATGGTGGAAAGAACTCTGGAAGGAGGCGGTGCGAGCG 299  
 QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
 DB 300 GTCTGACCACTGACTACCAAGCCCTCGTGGGCTCAATATACATTTCCGGTGAACCTG 359  
 QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
 DB 360 ATATTCCCTAGATGCCAAAGAGAGATGCCAATGGCAATGGAATGATCTATGAGAAAGATGC 419  
 QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAsp 140  
 DB 420 AGAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACATGGTCAGAG 479  
 QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyHisLeuArgPheProAspGlyLys 160  
 DB 480 GACAGTGACGGGGAATGGACCGCAAGCCATCATCAACGCTCTTCCCTGATGGGAA 539

QY 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
 DB 540 CCTTTTCTCCACCACCCCGGATGGAAAGATGGAATTTTCATCTAGCTCTTCCACACACTT 599  
 QY 181 GlyGlnTyrPheGlnLysGlnCysSerAlaArgValSerIleAsnThrValAsn 200  
 DB 600 GGTCAATATTTCCAGAAATTTGGGACGATGTTCACTGAGAGTTTCTGTGAACACACCAAT 659  
 QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220  
 DB 660 GTTACACTTGGGCTCAACTCATGGAAGTGAAGTGTCTGAGAGACATGAGCGGCATAT 719  
 QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
 DB 720 GTTCCCATCGCACAAAGTGAAGATGTTGCTGTAACAGATCAGATTCCTGTGTTGCTG 779  
 QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
 DB 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGAGCAACCTTCCC-AAAGATCTCCC 838  
 QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
 DB 839 ATTATGTTTGTGCTGCTGATTCATGATCCTAGCCACTTCTCAATATTATTCTACCAATTAAC 898  
 QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
 DB 899 TACAAGTGGAGCTTCGGGGATAACTGCGCTGTGTTGTTTCCACCACATCATCTGTGAAT 958  
 QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320  
 DB 959 CACAGTATGTGCTCAATGGAACTTTCAGCTTAACTCAGCTGTAAGAGCTGCAGACCA 1018  
 QY 321 GlyProCysProSerProThrProSerProSerSerSerThrSerProSerProAlaSer 340  
 DB 1019 GGACCTTGTCCGCCACCCACCCACCCAGACCT----- 1054  
 QY 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyrLysSer 360  
 DB 1055 -----TCAAAACCCACCCCTTCTTTAGGACCTGCTGTGACACCC 1096  
 QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380  
 DB 1097 CTGGAGCTGTAGTAGTATCTCTGATGANAATGCCAGATTAACACATATGCCACATTTCAA 1156  
 QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400  
 DB 1157 GCAACCATCACAAATGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTC 1216  
 QY 401 ProIleProThrLeuGlnProAsnSerLeuMetAspPheIleValThrCysLysGly 420  
 DB 1217 CTGATGCGCGGTGCCATGCGCTGAAAGCTCCCTAATAGACTTGTGCTGACCTGCCAAGG 1276  
 QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440  
 DB 1277 AGCATTTCCACGGAGGCTGTACCATCATTTCTGACCCACCTCGGAGATCCACCCAGAAC 1336  
 QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460  
 DB 1337 ACAATCTCGACGCCCTGTGGATGGATGAGATGCTGTGCTGAGCTGTGAGACGAACTTC 1396  
 QY 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480  
 DB 1397 AATCGGTCTGGAGCGTACTGTGTGAACCTCACCTGGGGATGACACACAGCTGCGCTCTC 1456  
 QY 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500  
 DB 1457 ACGACACCCCTGATTTCTGTTCTCTGACAGAGACCCAGCCCTCGCCTTTAAGGATGGCAAC 1516  
 QY 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
 DB 1517 AGTGCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGCTACATGTGATCTCCCTCTGTGG 1576  
 QY 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnGlyThrArgAsnValValLysGly 540



```
QY 396 GlnValAlaAspValProThrLeuGlnProAspAsnSerLeuMetAspPheile 415
DB 1519 CAGGCT-----GTGGCGTCCGGTGGGGGAT-----GCATTTAGCTGACT 1560

QY 416 ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435
DB 1561 GTGTCCTGCCAAGGGGGTGGCCCAAGGAGCTGCGATGGAGATCTCATCCAGGTGC 1620

QY 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuSer 455
DB 1621 CAGCCCGCTGCCAGCGGCTGTGCCAGCGCTGTCTACCCAGCCAGCCAGCCAGCTGGTT 1680

QY 456 ValArgArgAlaPhe---AsnGlySerGlyThrThrCysValAsnAspPheThrLeuGlyAsp 474
DB 1681 CTGCACCAAGTACTAGAGGGTGGCTCGGGGACATAGCTGCCCTCAATGTCTCTGGCTGAT 1740

QY 475 AspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGly--- 493
DB 1741 ACCAACAGCCTGGCAGTGGTGCACACCCAGCTTATCATGCTGTGGTCAAGAGCAGGCTT 1800

QY 494 SerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheVal 513
DB 1801 GGGCAGGTTCCGCTGATCGTGGGCATCTTG-----CTGCTGCTGATGGCT 1845

QY 514 ThrMetValThrIleLeuLeuThrLysLys 523
DB 1846 GTGGTCTTGCATCTGTATATATAGGCGC 1875

RESULT 13
US-09-862-260A-1
; Sequence 1, Application US/09862260A
; Patent No. US20020082217A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-MELANOMA COMPOUNDS
; FILE REFERENCE: 126881210200
; CURRENT APPLICATION NUMBER: US/09/862,260A
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/208,955
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/267,877
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ. ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO. 1
; LENGTH: 2130
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-862-260A-1

Alignment Scores:
Pred. No.: 1,56e-56 Length: 2130
Score: 627.50 Matches: 176
Percent Similarity: 41.85% Conservativeness: 96
Best local Similarity: 27.08% Mismatches: 205
Query Match: 20.50% Indels: 173
DB: 19 Gaps: 19

US-09-943-075A-2 (1-572) x US-09-862-260A-1 (1-2130)

QY 7 ValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGlnAlaAlaLys-----Arg 24
DB 46 CTCTTCATTGTTGGCTGTGATAGTGTCTTGGCTGTGGGGCTGACAAAAGTACCCAGA 105

QY 25 PheArgAspValLeuGlyHisGluInTyProAspHisMetArgGluAsnGlnLeu 44
DB 106 AACCAAGGACTGGCTTGGT-----GTCTCAAGGCAACTC 138

QY 45 ArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyProValTrpArg 64
DB 139 ACA-----ACCAAGCTGGACAGCAGCTGTATCCAGAGTGGACA--- 180

QY 65 GlyGluGlyArgTrpLysAspSerTrpGluGlyArgValGlnAlaAlaLeuThrSer 84
```

```
DB 181 ---GAAGCCCAAGAGACTGACTGCTGGAGAGGTGGTCAAGTGTCCCTCAAGGTGAGTAAT 237
QY 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeuValPheProArg 104
DB 238 GATGGGCTACACTGATTGGTCAAAATGCTCTTCTCTATTGGCTTGAATCTCCCTGGA 297
QY 105 CysGlnLysGluAspAlaAsnGlyAsnIleValTyrgluArgAsn---CysArgSerAsp 123
DB 298 AGCAAAAGGTATTGGCAGATGGGAGGTATTCTGGGTCAACAATACCATCATCAATGGG 357
QY 124 LeuGluLeuAlaSerAspProTyrgluValTyrgluValAsnTrpThrGlyAlaAspGluAsp 143
DB 358 AGCCAGGTGGGGAGGAGCAGCCAGTGTATCCCAAGAACT-----GACGATGCC--- 408
QY 144 TrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPhePro 163
DB 409 -----TGCACTCTCCCTGATGGTGGGACCTTGGCCA 438
QY 164 ArgProHisGlyArgLysLysTrpAsnPheValTyrgluPheHisThrLeuGlyGlnTy 183
DB 439 TCTGGCTCTTGGTCTCAGAAGAGAGCTTGTATTGTCTGGAAGACCTGGGGCAATAC 498
QY 184 PheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsnLeuThrVal 203
DB 499 TGGCAAGTCTAGGGGGCCCGAGTGTCTGGCTGAGCATTTGGCAGGAGGCGCATGCTG 558
QY 204 GlyProGlnValMetGluValIleValPheArgArgHisGly---ArgAlaTyriLePro 222
DB 559 GGCACACACACCATGGAAAGTGTCTACCATCGCGGGGATCCCGGAGTATGTGCT 618
QY 223 IleSerLysValLysAspValTyrgluValIleThrAspGlnIlePheValThrMet 242
DB 619 CTTCTCATTTCCAGCTCAGCTTCACCATTTACTGACAGGTGCCCTTCTCGTGAGCGGTG 678
QY 243 TyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhe 262
DB 679 TCCAGAGTTCGGGGCTTTGGATGGAGGGAACAAGCACTTCTCTGAGAAATCAGCCCT 738
QY 263 PheAspValLeuIleHisAspProSerHisPheLeuAsnTyrgluSerAlaIleSerTyrg 282
DB 739 TTTTCCCTCCAGCTCCATGACCCAGTGGCTATCTGCTGAGCTGACCTCTCTACACC 798
QY 283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302
DB 799 TGGGACTTTGGAGACAGTGTGGAACCCCTGATCTCTCGGGCAGCTTGTGGTCACTCAT 858
QY 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro----- 320
DB 859 TACTTGGAGCCTGGCCCAAGTCACTGGCCAGGTGGTCTCTGAGGCTGCCATTCCTCTC 918
QY 321 -----GlyProCysPro----- 324
DB 919 TCCTGTGGCTCTCTCCCAAGTCTCCAGGACCAACAGATGGGACAGGCCAATGCAGNGCC 978
QY 324 ----- 324
DB 979 CCTPACACCAAGCTGGCCCAAGTCCCTACTACAGAAAGTTGGTGTACTACACCTGGTCA 1038
QY 325 SerProThrProSerProSerSerSerThr----- 334
DB 1039 GCGCAACTGCAGAGCCCTCTGGAACCAACATCTGTGCAGTGCACCAACCACTGAAGTCATA 1098
QY 335 SerProSerProAlaSerSerProSerProThrLeuSerThrProSerSerSerLeuMet 354
DB 1099 AGCACTGCACTGTGTCAGATGCCAAGTGCAGAGAGCAGAGTATGACACCTGAGAAGTG 1158
QY 355 Pro-----ThrGly 357
DB 1159 CCAGTTTCAGAGGTGATGGTACCAACTGGCAGAGATGTCACTCCAGAGCTACAGGT 1218
QY 358 TyrLysSerMetGluLeuSerAspIle----- 366
```

```
Db 1219 ATGACACCTGCAGAGGTATCAATTGTGGTCTTCTTGGAAACACACAGCTGCACAGGTAACA 1278
Qy 366 -----
Db 1279 ACTACAGAGTGGGTGGAGACACACAGCTAGACAGCTACCTATCCCTGAGCCTGAAGGTCCA 1338
Qy 367 -----SerAsnGlu-----
Db 1339 GATCCAGCTCAATCATCTCTACGGAAAGATTACAGGTTCCCTGGGCCCCCTCCTGGAT 1398
Qy 370 -----AsnCysArgIleAsnArg 375
Db 1399 GGTACAGCCACTTAAGCTGGTGTGAAGAGACAAGTCCCCCTGGATTGTGTTCTGATATCGA 1458
Qy 376 TyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle 395
Db 1459 TATGTTCTTCTCCGTCACCTGGACATGTCACAGGGTATTGAAAGTGCCGAGATCCTG 1518
Qy 396 GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
Db 1519 CAGGCT-----GTGCGCTCCGTGAGGGGAT-----GCATTGAGCTGACT 1560
Qy 416 ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435
Db 1561 GTGCTCTGCCAAGCGGGCTGCCAAGGAAGCTGATGAGATCTCATGCCAGGGTGC 1620
Qy 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuSer 455
Db 1621 CAGCCCCCTGCCAGCGGCTGTGCCAGCTGTGCTACCCAGCCAGCCAGCTGCCAGCTGGTT 1680
Qy 456 ValArgArgAlaPhe---AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAsp 474
Db 1681 CTGCACACAGATCTGAAGGTGCTCGGGACATCTGCTCAATGTCTCTGCTGCTGAT 1740
Qy 475 AspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGly--- 493
Db 1741 ACCAAGACCTGGCAGTGTGCACACCCAGCTTATCATGCTGTGTCAGAAAGAGCGCTT 1800
Qy 494 SerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheVal 513
Db 1801 GGGCAGGTTCGGCTGATCGTGGGCATCTTG-----CTGGTGTGATGGCT 1845
Qy 514 ThrMetValThrIleLeuLeuTyrLysLys 523
Db 1846 GTGGTCTTGCATCTCTGATATATAGCGC 1875

RESULT 14
US-10-047-539-3
; Sequence 3, Application US/10047539
; Patent No. US20020177547A1
; GENERAL INFORMATION:
; APPLICANT: MOLLING, KARIN
; APPLICANT: PAVLOVIC, JOVAN
; APPLICANT: NAWRATH, MICHAEL
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS FOR TREATING OR PREVENTING
; FILE OF INVENTION: CANCER
; FILE REFERENCE: VOS-27
; CURRENT APPLICATION NUMBER: US/10/047,539
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: EP 01 10 0914.9
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2131
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)..(2018)
US-10-047-539-3
```

```
Pred. No.: 5,99e-56 Length: 2131
Score: 622.00 Matches: 177
Percent Similarity: 41.55% Conservative: 96
Best Local Similarity: 26.94% Mismatches: 204
Query Match: 20.32% Indels: 180
DB: Gaps: 20

US-09-943-075a-2 (1-572) x US-10-047-539-3 (1-2131)
Qy 7 ValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGlnAlaAlaLys-----Arg 24
Db 36 CTTCTTCATTGGCTGTGATAGTGTCTTCGGCTGTGGGGCTACAAAGTAGTACCCAGA 95
Qy 25 PheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGluAsnGlnLeu 44
Db 96 AACCAAGACTGGCTTGGT-----GTCTCAAGGCAACTC 128
Qy 45 ArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrProValTrpArgArg 64
Db 129 AGA-----ACCAAGCTTGGACAGCGCAGCTGTATCCAGAGTGGACA--- 170
Qy 65 GlyGluGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAlaLeuThrSer 84
Db 171 ---GAAGCCCAAGACTGTGCTGTGAGAGGTGTCAAGTGTCCCTCAAGGTCAAGTAAT 227
Qy 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeuValPheProArg 104
Db 228 GATGGCCCTACACTGATGTTGTCGAATGCTCTCTCTATTGCTTGAACCTTCCCTGGA 287
Qy 105 CysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsn---CysArgSerAsp 123
Db 288 AGCCAAAAGGTATTGCCAGATGGCAGGTATGCTGGGTCAACAATACCATCATCAATGGG 347
Qy 124 LeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAspAspGluAsp 143
Db 348 AGCCAGGTGTGGGAGGACAGCCAGTGTATCCCAAGAACT-----GACGATGCC--- 398
Qy 144 TrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPhePro 163
Db 399 -----TGCATCTTCCCTGATGGTGGAGCTTGGCCCA 428
Qy 164 ArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeuGlnTyr 183
Db 429 TCTGCTCTTGGTCTCAGAAGAGAGCTTGTATTATGCTGGAAGACCTTGGGGCCAAATAC 488
Qy 184 PheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsnLeuThrVal 203
Db 489 TGGCAAGTCTTAGGGGGCCCAAGTGTCTGGCTGAGCATTTGGGACAGGCGCAATGCTG 548
Qy 204 GlyProGlnValMetGluValIleValPheArgArgHisGly---ArgAlaTyrIlePro 222
Db 549 GGCACACACACCATGGAAGTGAAGTGTCTACCATCGCGGGGATCCCGGAGCTATGCTCT 608
Qy 223 IleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheValPheMet 242
Db 609 CTGCTCATTCACCTCAGCTTCACCATTCACCATTCACAGGTCCTTCTCCTGAGCGTG 668
Qy 243 TyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhe 262
Db 669 TCCCAAGTTCGGGCTTGGATGGAGGACACAGCACTCTCTGAGAAATCAGCTCTGACC 728
Qy 263 PheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLys 282
Db 729 TTTGCCCTCCAGCTCCATGACCTAGTGGCTATCTGGCTGAGCTGACCTCTCTACACC 788
Qy 283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302
Db 789 TGGGACTTTGGAGACAGTAGTGAACCCCTGATCTCTCGGGCAGCTGTGGTCTACTACTAT 848
Qy 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro----- 320
Db 849 TACCTGGAGCGTGGCCCGCAGTCACTGCCAGGTGCTCTGCAGGCTGCCATCTCTCTCACC 908

Alignment Scores:
```





Db 640 CTTTCTCCGTGAGTGTGTCCTCCAGCTACAGCCCTTGGACGGAGAGACCAAGCACTTCCCTG 699  
 Qy 257 ArgAspLeuProIlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyr 276  
 Db 700 AGAAATCATCTCTATCTTCCCTTCAGCTCCAGCCACCCAGTGTATTTGGCCGAG 759  
 Qy 277 SerAlaIleSerTyrTyrPheAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsn 296  
 Db 760 GCGACCTCTCTCATACATGGGACTTTGGAGATGGTACTGGACCCCTGTATCTCTCGGCA 819  
 Qy 297 HisThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGln 316  
 Db 820 CTGTGATGTCACCTACCTACCTGGCTGGCTCAGTCACCTGCCAGGTGGTACTGCAG 879  
 Qy 317 ThrAlaValPro-----GlyProCysProSerPro----- 326  
 Db 880 GTGCCATTTCTCTTTTCTGTGTCTCTCCCTCCAGCTCCCGGTACCCAGATGGCTAC 939  
 Qy 326 ----- 326  
 Db 940 ATCCCAACTGCAGAGCACCTGGAACACATCTAGGCAAGAACCACTACAAAAGTTGTG 999  
 Qy 327 -----ThrProSerProSerSerSerThrProSer----- 337  
 Db 1000 GGTACTACACCTGGCCAGATGCCAATACACAGCCCTCTGGAACACACAGTGTACAAATG 1059  
 Qy 338 ProAlaSerProSerProThrLeuSer----- 347  
 Db 1060 CCAACCCAGAGGTCCACAGCTACTACATCTGAGCAGATGCTGACCTCAGCGGTCTAGAT 1119  
 Qy 348 -----ThrProSer-----ProSer 352  
 Db 1120 ACCACACTGGCAGAGGTGTCAACTACAGAGGGTACAGGTACCCACCCACAGGCCCTTCT 1179  
 Qy 353 -----LeuMetPro 355  
 Db 1180 GGAACACCGTTGCAGAACCAACACACAGAGGTCACAGTCCAGCCCATTTGTGCC 1239  
 Qy 356 ThrGlyTyrIleSerMetGluLeuSerAspIleSerAsnGlu----- 369  
 Db 1240 ACACAAAAGTTCTACAGGGTCCATTTAGCCCTCTACTGATGACACCGACACCAATAGTCTT 1299  
 Qy 370 -----AsnCysArgIleAsnArgTyrGlyTyrPheArgAlaThr 382  
 Db 1300 GTGACAGACAAGTTCCTCCCTGGACTGTCTCTATATGATATGGTCTTCTCCCTGCC 1359  
 Qy 383 IleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspValProIle 402  
 Db 1360 CTGGACATTTGTCAGGGTATTGAAAGTGTGAGATCTCTGAGGCT-----GTGCCATT 1413  
 Qy 403 ProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGlyAlaThr 422  
 Db 1414 -----AGTGAAGGGATGTCATTTGAGCTGACTGTGCTGCCAAGCGGGCTA 1461  
 Qy 423 ProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsnArgVal 442  
 Db 1462 CCAAGGAAGCCGTGTGACATTTTCATCAGGGTGCCAGGCCCTGCCAGAGGCTG 1521  
 Qy 443 CysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe-----Asn 461  
 Db 1522 TGCCAGTGTGTTCCACCGAGCCACAGTGCAGCTGGTCTTACACCAAGTGTGAAAGGT 1581  
 Qy 462 GlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThr 481  
 Db 1582 GGCTCAGGCACATATTGCTCAATGTGCTTTGGCTGACGCCAACACGCTGGCAGTGGCC 1641  
 Qy 482 SerAlaLeuIleSerIleProGlyLysAspLeuGly-----SerProLeuArgThr 498  
 Db 1642 AGCCACCAACTTGTGTTCTCGTGGTCAAGCGGTGGCTGGCAGGCTCCCTTG----- 1695  
 Qy 499 ValAsnGlyValIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIle 518  
 Db 1696 -----CTTCTAGGTATCTGTGTTGGTGTGGCTGTGGTCTTGCATCTCTG 1743

Qy 519 LeuLeuTyrTyrLysHisLysThrTyrIleHisProIleGlyAsnCysThrArgAsnValVal 538  
 Db 1744 ATACTAGGCATAGACTTAAGAAGCAGGGCTCAGTTTCCCAAAATGCCACATG----- 1794  
 Qy 539 LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAsp 558  
 Db 1795 -----GTAGCACTCAGTGTGGCTGCTCGGCTCGCTCGGCTTCGGCGCCCGGGCTT 1842  
 Qy 559 ArgGluLysAspProLeuLeu 565  
 Db 1843 GGAGAAAACAGCCCGCTCCTC 1863

Search completed: April 9, 2003, 16:00:55  
 Job time : 181 secs